Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US05/010913

International filing date: 31 March 2005 (31.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/559,275
Filing date: 01 April 2004 (01.04.2004)

Date of receipt at the International Bureau: 12 August 2005 (12.08.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1352384

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME;

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

August 02, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/559,275*
FILING DATE: *April 01, 2004*
RELATED PCT APPLICATION NUMBER: *PCT/US05/10913*

Certified by

[signature] W. Dudas

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

Abstract of the Disclosure

Provided herein are methods for identifying a risk of osteoarthritis in a subject, reagents and kits for carrying out the methods, methods for identifying candidate therapeutics for treating osteoarthritis, and therapeutic and preventative methods applicable to osteoarthritis. These embodiments are based upon an analysis of polymorphic variations in nucleotide sequences within the human genome.

Application Data Sheet

Application Information

Application Type::	Provisional
Subject Matter::	Utility
Suggested Group Art Unit::	Not Yet Assigned
CD-ROM or CD-R?::	None
Sequence submission?::	None
Computer Readable Form (CRF)?::	No
Title::	METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF
Attorney Docket Number::	524593009000
Request for Early Publication?::	No
Request for Non-Publication?::	No
Total Drawing Sheets?::	3
Small Entity?::	Yes
Petition included?::	No
Secrecy Order in Parent Appl.?::	No

Applicant Information

Applicant Authority Type::	Inventor
Primary Citizenship Country::	US
Status::	Full Capacity
Given Name::	Steven
Family Name::	MAH
City of Residence::	San Diego
State or Province of Residence::	CA
Country of Residence::	US
Street of mailing address::	12820 Via Nieve #74
City of mailing address::	San Diego
State or Province of mailing address::	CA

Initial 04/01/04

Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: Germany
Status:: Full Capacity
Given Name:: Andreas
Family Name:: BRAUN
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 3935 Lago Di Grata Circle
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: Germany
Status:: Full Capacity
Given Name:: Stefan
Middle Name:: M.
Family Name:: KAMMERER
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 3825 Elijah Court, Unit 334
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: US
Status:: Full Capacity

Initial 04/01/04

Given Name::   Matthew

Middle Name::   Roberts

Family Name::   NELSON

City of Residence::   San Marcos

State or Province of Residence::   CA

Country of Residence::   US

Street of mailing address::   1250 Calle Prospero

City of mailing address::   San Marcos

State or Province of mailing address::   CA

Postal or Zip Code of mailing address::   92069

Applicant Authority Type::   Inventor

Primary Citizenship Country::   Sweden

Status::   Full Capacity

Given Name::   Rikard

Middle Name::   Henry

Family Name::   RENELAND

City of Residence::   San Diego

State or Province of Residence::   CA

Country of Residence::   US

Street of mailing address::   7555 Charmant Drive, #1114

City of mailing address::   San Diego

State or Province of mailing address::   CA

Postal or Zip Code of mailing address::   92122

Applicant Authority Type::   Inventor

Primary Citizenship Country::   United Kingdom

Status::   Full Capacity

Given Name::   Maria

Middle Name::   L.

Family Name::   LANGDOWN

City of Residence::   San Diego

State or Province of Residence::   CA

Country of Residence::   US

Street of mailing address::   3701 Yosemite Street

City of mailing address::   San Diego

State or Province of mailing address::   CA

Postal or Zip Code of mailing address::   92109

Correspondence Information

Correspondence Customer Number::   25225

Representative Information

Representative Customer Number::   25225

Initial 04/01/04

What is claimed is:

1.   A method for identifying a subject at risk of osteoarthritis, which comprises detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the one or more polymorphic variations are detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-4;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(d) a fragment of a nucleotide sequence of (a), (b), or (c);

whereby the presence of the polymorphic variation is indicative of the subject being at risk of osteoarthritis.

2.   The method of claim 1, which further comprises obtaining the nucleic acid sample from the subject.

3.   The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 102570000 to 102583000 in human genomic DNA.

4.   The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions selected from the group consisting of 207, 6019, 6414, 7341, 10984, 12351, 13335, 16584, 16737, 23897, 24057, 25145, 25300, 26262, 26312, 26589, 27302, 27358, 27451, 27552, 30731, 32085, 32139, 33184, 42382, 42569, 44823, 45217, 45548, 45601, 45722, 45967, 47367, 47642, 48126, 49218, 49274, 49433, 49610, 51282, 51466, 53757, 53960, 54031, 54574, 55679, 56100, 56182, 59817, 60533, 60656, 72209, 72778, 74293, 77335, 78029, 78374, 78421, 78434, 79174, 79397, 79562, 79700, 79730, 79904, 79920, 79938, 79972, 80125, 80368, 83484, 85536, 85829, 86425, 88083, 88770, 90622, 90924, 91634, 92029, 95152, 95348, 96145, 96793, 97015, 97064, 97711, 97855 and 98708.

5.   The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 1 selected from the group consisting of 6414, 51282, 54574, 78374, 92029 and 96793.

6. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in linkage disequilibrium with one or more positions in claim 3, 4 or 5.

7. The method of claim 1, wherein detecting the presence or absence of the one or more polymorphic variations comprises:

hybridizing an oligonucleotide to the nucleic acid sample, wherein the oligonucleotide is complementary to a nucleotide sequence in the nucleic acid and hybridizes to a region adjacent to the polymorphic variation;

extending the oligonucleotide in the presence of one or more nucleotides, yielding extension products; and

detecting the presence or absence of a polymorphic variation in the extension products.

8. The method of claim 1, wherein the subject is a human.

9. The method of claim 8, wherein the subject is a human female.

10. The method of claim 8, wherein the subject is a human male.

11. A method for identifying a polymorphic variation associated with osteoarthritis proximal to an incident polymorphic variation associated with osteoarthritis, which comprises:

identifying a polymorphic variation proximal to the incident polymorphic variation associated with osteoarthritis, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

(a) a nuclcotide sequence in SEQ ID NO: 1-4;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation;

determining the presence or absence of an association of the proximal polymorphic variant with osteoarthritis.

12. The method of claim 11, wherein the incident polymorphic variation is at one or more positions in claim 3, 4 or 5.

13. The method of claim 11, wherein the proximal polymorphic variation is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the incident polymorphic variation.

14. The method of claim 11, which further comprises determining whether the proximal polymorphic variation is in linkage disequilibrium with the incident polymorphic variation.

15. The method of claim 11, which further comprises identifying a second polymorphic variation proximal to the identified proximal polymorphic variation associated with osteoarthritis and determining if the second proximal polymorphic variation is associated with osteoarthritis.

16. The method of claim 15, wherein the second proximal polymorphic variant is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the proximal polymorphic variation associated with osteoarthritis.

17. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-4;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

(e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d);

wherein the nucleotide sequence comprises a polymorphic variation associated with osteoarthritis selected from the group consisting of an adenine at position 6414, an adenine at positoin 51282, a cytosine at position 54574, a thymine at position 92029 and an adenine at position 96793.

18. An oligonucleotide comprising a nucleotide sequence complementary to a portion of the nucleotide sequence of (a), (b), (c), or (d) in claim 17, wherein the 3' end of the oligonucleotide is adjacent to a polymorphic variation associated with osteoarthritis.

19. A microarray comprising an isolated nucleic acid of claim 17 linked to a solid support.

116

20. An isolated polypeptide encoded by the isolated nucleic acid sequence of claim 17.

21. A method for identifying a candidate therapeutic for treating osteoarthritis, which comprises:

(a) introducing a test molecule to a system which comprises a nucleic acid comprising a nucleotide sequence selected from the group consisting of:

(i)   a nucleotide sequence in SEQ ID NO: 1-4;

(ii)   a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(iii)   a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(iv)   a fragment of a nucleotide sequence of (a), (b), or (c); or

introducing a test molecule to a system which comprises a protein encoded by a nucleotide sequence of (i), (ii), (iii), or (iv); and

(b) determining the presence or absence of an interaction between the test molecule and the nucleic acid or protein,

whereby the presence of an interaction between the test molecule and the nucleic acid or protein identifies the test molecule as a candidate therapeutic for treating osteoarthritis.

22. The method of claim 21, wherein the system is an animal.

23. The method of claim 21, wherein the system is a cell.

24. The method of claim 21, wherein the nucleotide sequence comprises one or more polymorphic variations associated with osteoarthritis.

25. The method of claim 24, wherein the one or more polymorphic variations associated with osteoarthritis are at one or more positions in claim 3, 4 or 5.

26. A method for treating osteoarthritis in a subject, which comprises contacting one or more cells of a subject in need thereof with a nucleic acid, wherein the nucleic acid comprises a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-4;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-4;

117

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(d) a fragment of a nucleotide sequence of (a), (b), or (c); and

(e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d);

whereby contacting the one or more cells of the subject with the nucleic acid treats the osteoarthritis in the subject.

27. The method of claim 26, wherein the nucleic acid is RNA or PNA.

28. The method of claim 27, wherein the nucleic acid is duplex RNA.

29. A method for treating osteoarthritis in a subject, which comprises contacting one or more cells of a subject in need thereof with a protein, wherein the protein is encoded by a nucleotide sequence which comprises a polynucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-4;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(d) a fragment of a nucleotide sequence of (a), (b), or (c);

whereby contacting the one or more cells of the subject with the protein treats the osteoarthritis in the subject.

30. A method for treating osteoarthritis in a subject, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the one or more polymorphic variation are detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-4;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

administering an osteoarthritis treatment to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

31. The method of claim 30, wherein the one or more polymorphic variations are detected at one or more positions in claim 3, 4 or 5.

32. The method of claim 30, wherein the treatment is selected from the group consisting of administering a corticosteroid, a corticosteroid, a nonsteroidal anti-inflammatory drug (NSAID), a cyclooxygenase-2 (COX-2) inhibitor, an antibody, a glucocorticoid, hyaluronic acid, chondrotin sulfate, glucosamine or acetaminophen; prescribing a heat/cold regimen or a joint protection regimen; performing joint surgery; prescribing a weight control regimen; and combinations of the foregoing.

33. A method for detecting or preventing osteoarthritis in a subject, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-4;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

administering an osteoarthritis prevention or detection procedure to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

34. The method of claim 33, wherein the one or more polymorphic variations are detected at one or more positions in claim 3, 4 or 5.

35. The method of claim 33, wherein the osteoarthritis prevention is selected from the group consisting of administering a corticosteroid, a nonsteroidal anti-inflammatory drug (NSAID), a cyclooxygenase-2 (COX-2) inhibitor, an antibody, a glucocorticoid, hyaluronic acid, chondrotin sulfate, glucosamine or acetaminophen; prescribing a heat/cold regimen or a joint protection regimen; performing joint surgery; prescribing a weight control regimen; and combinations of the foregoing.

36. A method of targeting information for preventing or treating osteoarthritis to a subject in need thereof, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-4;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-4;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

directing information for preventing or treating osteoarthritis to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

37. The method of claim 36, wherein the one or more polymorphic variations are detected at one or more positions in claim 3, 4 or 5.

38. A composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and an antibody that specifically binds to a protein, polypeptide or peptide encoded by a nucleotide sequence identical to or 90% or more identical to a nucleotide sequence in SEQ ID NO: 1-4.

39. A composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and a RNA, DNA, PNA or ribozyme molecule comprising a nucleotide sequence identical to or 90% or more identical to a portion of a nucleotide sequence in SEQ ID NO: 1-4.

40. The composition of claim 39, wherein the RNA molecule is a short inhibitory RNA molecule.
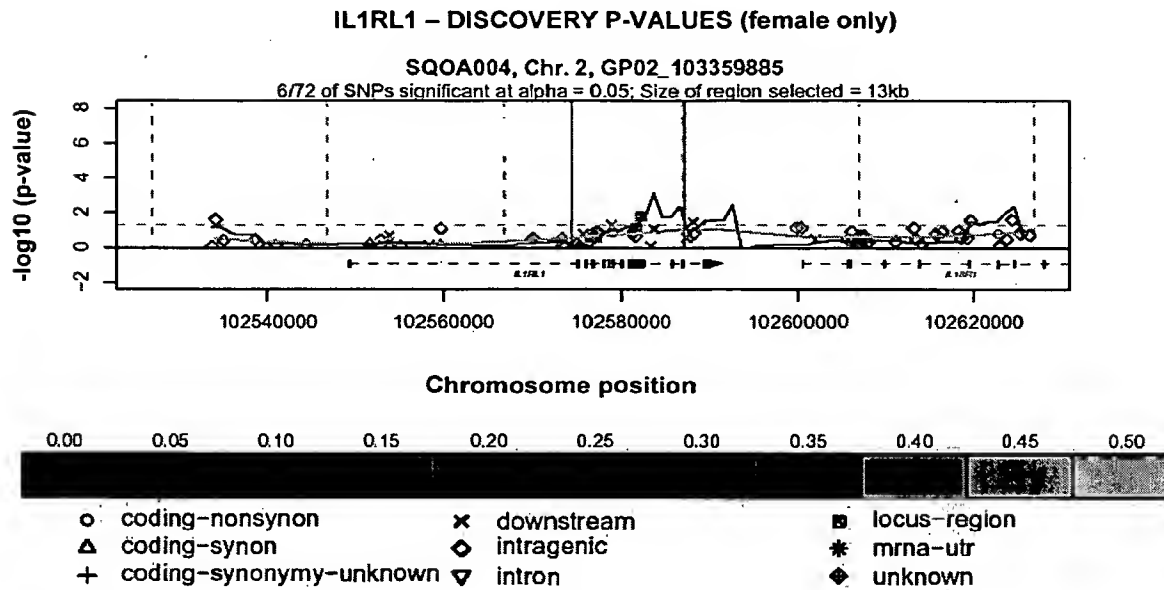
FIGURE 1

IL1RL1 – DISCOVERY P-VALUES (female only)

FIGURE 2

IL1RL1 isoform expression (SW1353 monolayers)
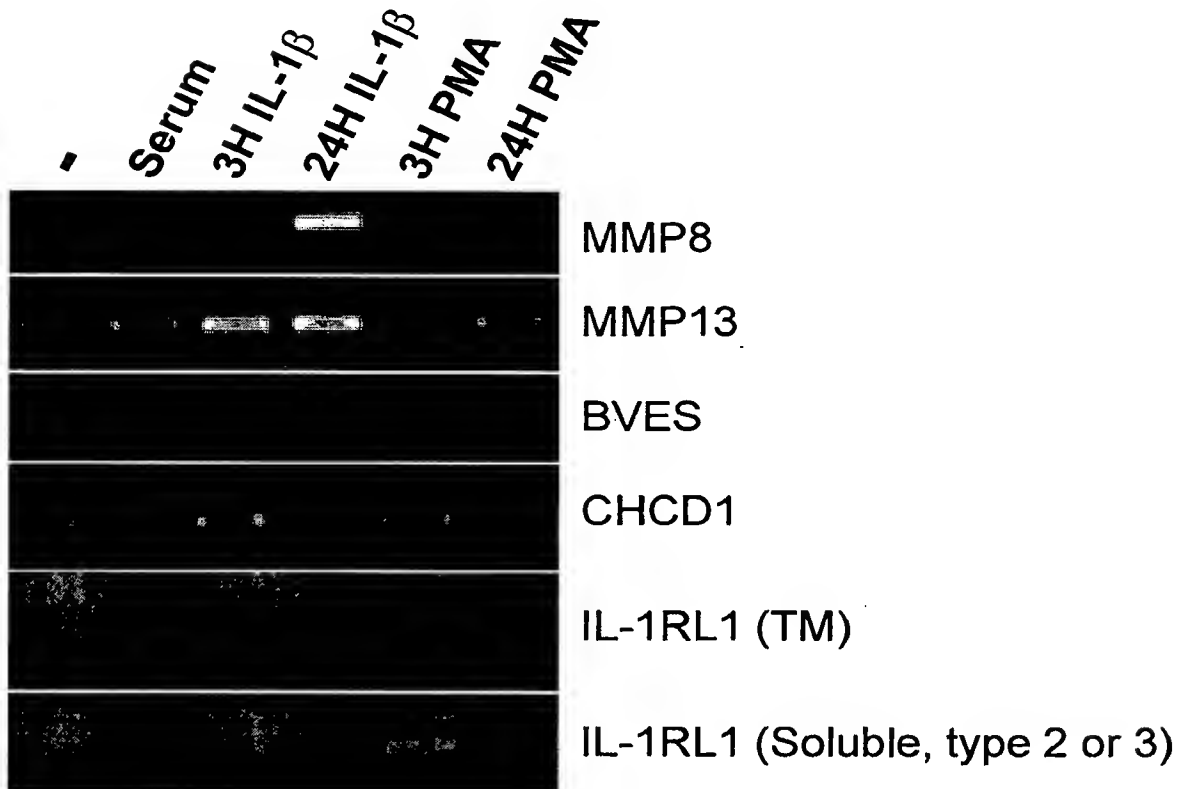
FIGURE 3A

IL1RL1 isoform expression (3-D alginate)

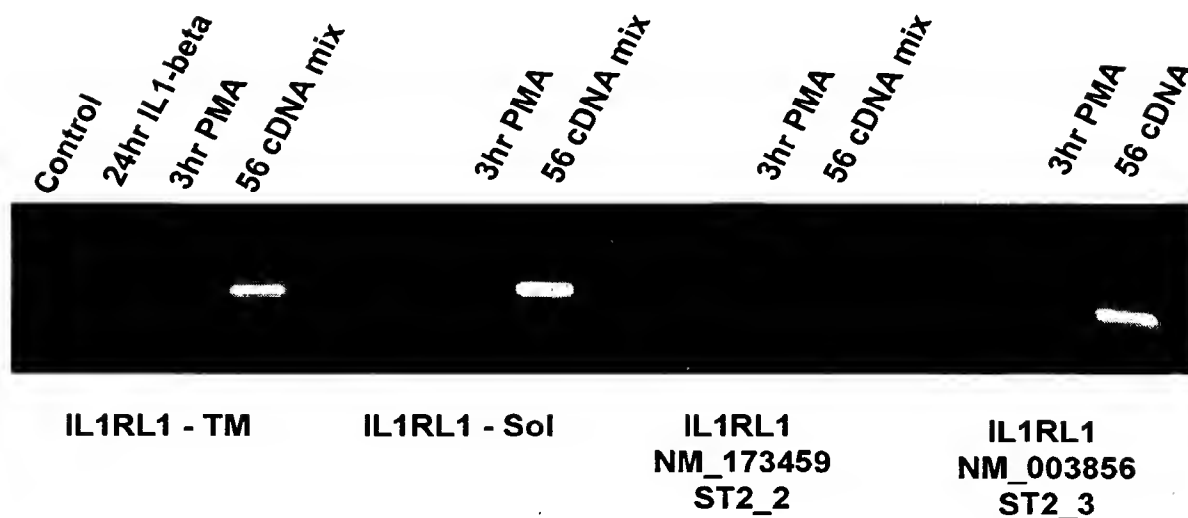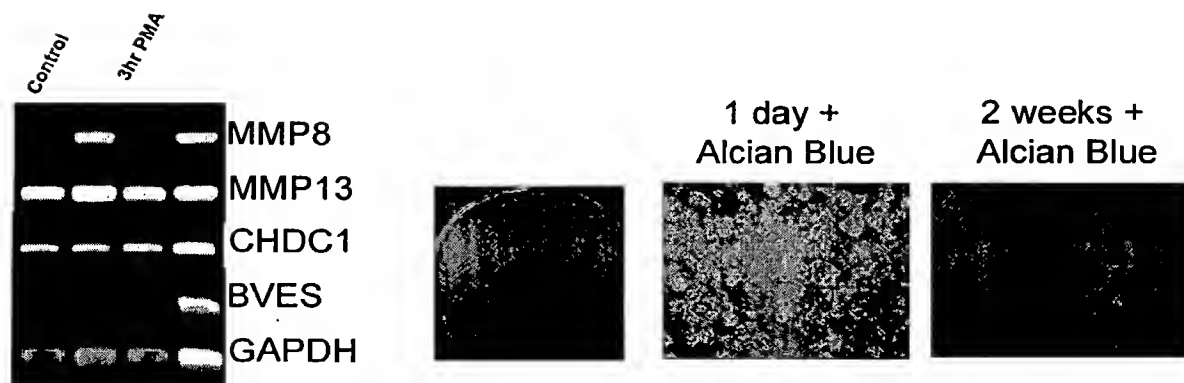| IL1RL1 - TM | IL1RL1 - Sol | IL1RL1 NM_173459 ST2_2 | IL1RL1 NM_003856 ST2_3 |

FIGURE 3B

IL1RL1 isoform expression (3-D alginate)

METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF

Field of the Invention

[0001] The invention relates to genetic methods for identifying risk of osteoarthritis and treatments that specifically target such diseases.

Background

[0002] Osteoarthritis (OA) is a chronic disease usually affecting weight-bearing synovial joints. There are approximately 20 million Americans affected by OA and it is the leading cause of disability in the United States. In addition to extensive human suffering, OA also accounts for nearly all knee replacements and more than half of all hip replacements in the United States. Despite its prevalence, OA is poorly understood and there are few treatments available besides anti-inflammatory drugs and joint replacement.

[0003] Most commonly affecting middle-aged and older people, OA can range from very mild to very severe. It affects hands and weight-bearing joints such as knees, hips, feet and the back. Knee OA can be as disabling as any cardiovascular disease except stroke.

[0004] OA is characterized by the breakdown of cartilage in joints. Cartilage in joints cushions the ends of bones, and cartilage breakdown causes bones to rub against each other, causing pain and loss of movement. Type II collagen is the main component of cartilage, comprising 15-25% of the wet weight, approximately half the dry weight, and representing 90-95% of the total collagen content in the tissue. It forms fibrils that endow cartilage with tensile strength (Mayne, R. Arthritis Rhuem. 32:241-246 (1989)).

Summary

[0005] It has been discovered that certain polymorphic variations in human genomic DNA are associated with osteoarthritis. In particular, polymorphic variants in a locus containing a *IL1RL1* region in human genomic DNA have been associated with risk of osteoarthritis.

[0006] Thus, featured herein are methods for identifying a subject at risk of osteoarthritis and/or a risk of osteoarthritis in a subject, which comprise detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in or around the loci described herein in a human nucleic acid sample. In an embodiment, two or more polymorphic variations are detected in two or more regions of which one is the *IL1RL1* region. In certain embodiments, 3 or more, or 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more polymorphic variants are detected.

[0007] Also featured are nucleic acids that include one or more polymorphic variations associated with occurrence of osteoarthritis, as well as polypeptides encoded by these nucleic acids. In addition, provided are methods for identifying candidate therapeutic molecules for treating osteoarthritis, as well as methods for treating osteoarthritis in a subject by identifying a subject at risk of osteoarthritis and treating the subject with a suitable prophylactic, treatment or therapeutic molecule.

[0008] Also provided are compositions comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and/or a *IL1RL1* nucleic acid, with a RNAi, siRNA, antisense DNA or RNA, or ribozyme nucleic acid designed from a *IL1RL1* nucleotide sequence. In an embodiment, the RNAi, siRNA, antisense DNA or RNA, or ribozyme nucleic acid is designed from a *IL1RL1* nucleotide sequence that includes one or more polymorphic variations associated with osteoarthritis, and in some instances, specifically interacts with such a nucleotide sequence. Further, provided are arrays of nucleic acids bound to a solid surface, in which one or more nucleic acid molecules of the array have a *IL1RL1* nucleotide sequence, or a fragment or substantially identical nucleic acid thereof, or a complementary nucleic acid of the foregoing. Featured also are compositions comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and/or a *IL1RL1* polypeptide, with an antibody that specifically binds to the polypeptide. Thus, featured is an antibody that specifically binds to an epitope in the polypeptide that includes an amino acid encoded by a polymorphic site associated with osteoarthritis. In certain embodiments, the antibody specifically binds to an epitope comprising a glutamate or alanine encoded by rs1041973 (e.g., an antibody that binds to an epitope comprising an alanine at position 78 in an *IL1RL1* polypeptide) An alanine at position 78 is associated with increased risk of osteoarthitis.

Brief Description of the Drawings

[0009] Figure 1 shows proximal SNPs in a *IL1RL1* region in genomic DNA. The position of each SNP in the chromosome is shown on the x-axis and the y-axis provides the negative logarithm of the p-value comparing the estimated allele to that of the control group. Also shown in the figure are exons and introns of the gene in the approximate chromosomal positions.

[0010] Figure 2 shows expression profiling results for *IL1RL1*.

[0011] Figures 3A and 3B show *IL1RL1* expression modulation in a human chondrocyte cell line model.

Detailed Description

[0012] It has been discovered that a polymorphic variant in a locus containing a *IL1RL1* region is associated with occurrence of osteoarthritis in subjects. Thus, detecting genetic determinants associated with an increased risk of osteoarthritis occurrence can lead to early identification of a predisposition to osteoarthritis and early prescription of preventative measures. Also, associating a *IL1RL1* polymorphic

variant with osteoarthritis has provided new targets for screening molecules useful in treatments of osteoarthritis.

Osteoarthritis and Sample Selection

[0013] Osteoarthritis (OA), or degenerative joint disease, is one of the oldest and most common types of arthritis. It is characterized by the breakdown of the joint's cartilage. Cartilage is the part of the joint that cushions the ends of bones, and its breakdown causes bones to rub against each other, causing pain and loss of movement. Type II collagen is the main component of cartilage, comprising 15-25% of the wet weight, approximately half the dry weight, and representing 90-95% of the total collagen content in the tissue. It forms fibrils that endow cartilage with tensile strength (Mayne, R. Arthritis Rhuem. 32:241-246 (1989)).

[0014] Most commonly affecting middle-aged and older people, OA can range from very mild to very severe. It affects hands and weight-bearing joints such as knees, hips, feet and the back. Knee OA can be as disabling as any cardiovascular disease except stroke.

[0015] Osteoarthritis affects an estimated 20.7 million Americans, mostly after age 45, with women more commonly affected than men. Physicians make a diagnosis of OA based on a physical exam and history of symptoms. X-rays are used to confirm diagnosis. Most people over 60 reflect the disease on X-ray, and about one-third have actual symptoms.

[0016] There are many factors that can cause OA. Obesity may lead to osteoarthritis of the knees. In addition, people with joint injuries due to sports, work-related activity or accidents may be at increased risk of developing OA.

[0017] Genetics has a role in the development of OA. Some people may be born with defective cartilage or with slight defects in the way that joints fit together. As a person ages, these defects may cause early cartilage breakdown in the joint or the inability to repair damaged or deteriorated cartilage in the joint.

[0018] Inclusion or exclusion of samples for an osteoarthritis pool may be based upon the following criteria: ethnicity (e.g., samples derived from an individual characterized as Caucasian); parental ethnicity (e.g., samples derived from an individual of British paternal and maternal descent); relevant phenotype information for the individual (e.g., case samples derived from individuals diagnosed with specific knee osteoarthritis (OA) and were recruited from an OA knee replacement clinic). Control samples may be selected based on relevant phenotype information for the individual (e.g., derived from individuals free of OA at several sites (knee, hand, hip etc)); and no family history of OA and/or rheumatoid arthritis. Additional phenotype information collected for both cases and controls may include age of the individual, gender, family history of OA, diagnosis with osteoarthritis (joint location of OA, date of primary diagnosis, age of individual as of primary diagnosis), knee history (current symptoms,

3

any major knee injury, menisectomy, knee replacement surgery, age of surgery), HRT history, osteoporosis diagnosis.

[0019] Based in part upon selection criteria set forth above, individuals having osteoarthritis can be selected for genetic studies. Also, individuals having no history of osteoarthritis often are selected for genetic studies, as described hereafter.

Polymorphic Variants Associated with Osteoarthritis

[0020] A genetic analysis provided herein linked osteoarthritis with polymorphic variant nucleic acid sequences in the human genome. As used herein, the term "polymorphic site" refers to a region in a nucleic acid at which two or more alternative nucleotide sequences are observed in a significant number of nucleic acid samples from a population of individuals. A polymorphic site may be a nucleotide sequence of two or more nucleotides, an inserted nucleotide or nucleotide sequence, a deleted nucleotide or nucleotide sequence, or a microsatellite, for example. A polymorphic site that is two or more nucleotides in length may be 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more, 20 or more, 30 or more, 50 or more, 75 or more, 100 or more, 500 or more, or about 1000 nucleotides in length, where all or some of the nucleotide sequences differ within the region. A polymorphic site is often one nucleotide in length, which is referred to herein as a "single nucleotide polymorphism" or a "SNP."

[0021] Where there are two, three, or four alternative nucleotide sequences at a polymorphic site, each nucleotide sequence is referred to as a "polymorphic variant" or "nucleic acid variant." Where two polymorphic variants exist, for example, the polymorphic variant represented in a minority of samples from a population is sometimes referred to as a "minor allele" and the polymorphic variant that is more prevalently represented is sometimes referred to as a "major allele." Many organisms possess a copy of each chromosome (e.g., humans), and those individuals who possess two major alleles or two minor alleles are often referred to as being "homozygous" with respect to the polymorphism, and those individuals who possess one major allele and one minor allele are normally referred to as being "heterozygous" with respect to the polymorphism. Individuals who are homozygous with respect to one allele are sometimes predisposed to a different phenotype as compared to individuals who are heterozygous or homozygous with respect to another allele.

[0022] In genetic analysis that associate polymorphic variants with osteoarthritis, samples from individuals having osteoarthritis and individuals not having osteoarthritis often are allelotyped and/or genotyped. The term "allelotype" as used herein refers to a process for determining the allele frequency for a polymorphic variant in pooled DNA samples from cases and controls. By pooling DNA from each group, an allele frequency for each SNP in each group is calculated. These allele frequencies are then compared to one another. The term "genotyped" as used herein refers to a process for determining a

genotype of one or more individuals, where a "genotype" is a representation of one or more polymorphic variants in a population.

[0023] A genotype or polymorphic variant may be expressed in terms of a "haplotype," which as used herein refers to two or more polymorphic variants occurring within genomic DNA in a group of individuals within a population. For example, two SNPs may exist within a gene where each SNP position includes a cytosine variation and an adenine variation. Certain individuals in a population may carry one allele (heterozygous) or two alleles (homozygous) having the gene with a cytosine at each SNP position. As the two cytosines corresponding to each SNP in the gene travel together on one or both alleles in these individuals, the individuals can be characterized as having a cytosine/cytosine haplotype with respect to the two SNPs in the gene.

[0024] As used herein, the term "phenotype" refers to a trait which can be compared between individuals, such as presence or absence of a condition, a visually observable difference in appearance between individuals, metabolic variations, physiological variations, variations in the function of biological molecules, and the like. An example of a phenotype is occurrence of osteoarthritis.

[0025] Researchers sometimes report a polymorphic variant in a database without determining whether the variant is represented in a significant fraction of a population. Because a subset of these reported polymorphic variants are not represented in a statistically significant portion of the population, some of them are sequencing errors and/or not biologically relevant. Thus, it is often not known whether a reported polymorphic variant is statistically significant or biologically relevant until the presence of the variant is detected in a population of individuals and the frequency of the variant is determined. Methods for detecting a polymorphic variant in a population are described herein, specifically in Example 2. A polymorphic variant is statistically significant and often biologically relevant if it is represented in 5% or more of a population, sometimes 10% or more, 15% or more, or 20% or more of a population, and often 25% or more, 30% or more, 35% or more, 40% or more, 45% or more, or 50% or more of a population.

[0026] A polymorphic variant may be detected on either or both strands of a double-stranded nucleic acid. Also, a polymorphic variant may be located within an intron or exon of a gene or within a portion of a regulatory region such as a promoter, a 5′ untranslated region (UTR), a 3′ UTR, and in DNA (e.g., genomic DNA (gDNA) and complementary DNA (cDNA)), RNA (e.g., mRNA, tRNA, and rRNA), or a polypeptide. Polymorphic variations may or may not result in detectable differences in gene expression, polypeptide structure, or polypeptide function.

[0027] It was determined that polymorphic variations associated with an increased risk of osteoarthritis existed in a *IL1RL1* region in SEQ ID NO: 1. In certain embodiments, a polymorphic variant at position rs1041973 in the human genome was associated with an increased risk of osteoarthritis, and in a specific embodiment, a cytosine at position rs1041973 was associated with an increased risk of osteoarthritis.

[0028] Polymorphic variants in and around the *IL1RL1* locus were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 1 selected from the group consisting of 207, 6019, 6414, 7341, 10984, 12351, 13335, 16584, 16737, 23897, 24057, 25145, 25300, 26262, 26312, 26589, 27302, 27358, 27451, 27552, 30731, 32085, 32139, 33184, 42382, 42569, 44823, 45217, 45548, 45601, 45722, 45967 ,47367, 47642, 48126, 49218, 49274, 49433, 49610, 51282, 51466, 53757, 53960, 54031, 54574, 55679, 56100, 56182, 59817, 60533, 60656, 72209, 72778, 74293, 77335, 78029, 78374, 78421, 78434, 79174, 79397, 79562, 79700, 79730, 79904, 79920, 79938, 79972, 80125, 80368, 83484, 85536, 85829, 86425, 88083, 88770, 90622, 90924, 91634, 92029, 95152, 95348, 96145, 96793, 97015, 97064, 97711, 97855 and 98708. Polymorphic variants at the following positions in SEQ ID NO: 1 in particular were associated with an increased risk of osteoarthritis: 6414, 51282, 54574, 78374, 92029 and 96793, where specific embodiments are directed to position 54574. In particular, the following polymorphic variants in SEQ ID NO: 1 were associated with risk of osteoarthritis: an adenine at position 6414, an adenine at positoin 51282, a cytosine at position 54574, a thymine at position 92029 and an adenine at position 96793.

[0029] Based in part upon analyses summarized in Figure 1, a region with significant association has been identified in a locus associated with osteoarthritis. Any polymorphic variants associated with osteoarthritis in a region of significant association can be utilized for embodiments described herein. For example, polymorphic variants in a region spanning chromosome positions 102570000 to 102583000 (approximately 13,000 nucleotides in length) in a *IL1RL1* locus have significant association (chromosome positions are within NCBI's Genome build 34).

Additional Polymorphic Variants Associated with Osteoarthritis

[0030] Also provided is a method for identifying polymorphic variants proximal to an incident, founder polymorphic variant associated with osteoarthritis. Thus, featured herein are methods for identifying a polymorphic variation associated with osteoarthritis that is proximal to an incident polymorphic variation associated with osteoarthritis, which comprises identifying a polymorphic variant proximal to the incident polymorphic variant associated with osteoarthritis, where the incident polymorphic variant is in a *IL1RL1* nucleotide sequence. The nucleotide sequence often comprises a polynucleotide sequence selected from the group consisting of (a) a polynucleotide sequence of SEQ ID NO: 1-4; (b) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence encoded by a polynucleotide sequence of SEQ ID NO: 1-4; and (c) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-4 or a polynucleotide sequence 90% or more identical to the polynucleotide sequence of SEQ ID NO: 1-4. The presence or absence of an association of the proximal polymorphic variant with osteoarthritis then is determined using a known

association method, such as a method described in the Examples hereafter. In an embodiment, the incident polymorphic variant is a polymorphic variant associated with osteoarthritis described herein. In another embodiment, the proximal polymorphic variant identified sometimes is a publicly disclosed polymorphic variant, which for example, sometimes is published in a publicly available database. In other embodiments, the polymorphic variant identified is not publicly disclosed and is discovered using a known method, including, but not limited to, sequencing a region surrounding the incident polymorphic variant in a group of nucleic samples. Thus, multiple polymorphic variants proximal to an incident polymorphic variant are associated with osteoarthritis using this method.

[0031] The proximal polymorphic variant often is identified in a region surrounding the incident polymorphic variant. In certain embodiments, this surrounding region is about 50 kb flanking the first polymorphic variant (e.g. about 50 kb 5' of the first polymorphic variant and about 50 kb 3' of the first polymorphic variant), and the region sometimes is composed of shorter flanking sequences, such as flanking sequences of about 40 kb, about 30 kb, about 25 kb, about 20 kb, about 15 kb, about 10 kb, about 7 kb, about 5 kb, or about 2 kb 5' and 3' of the incident polymorphic variant. In other embodiments, the region is composed of longer flanking sequences, such as flanking sequences of about 55 kb, about 60 kb, about 65 kb, about 70 kb, about 75 kb, about 80 kb, about 85 kb, about 90 kb, about 95 kb, or about 100 kb 5' and 3' of the incident polymorphic variant.

[0032] In certain embodiments, polymorphic variants associated with osteoarthritis are identified iteratively. For example, a first proximal polymorphic variant is associated with osteoarthritis using the methods described above and then another polymorphic variant proximal to the first proximal polymorphic variant is identified (e.g., publicly disclosed or discovered) and the presence or absence of an association of one or more other polymorphic variants proximal to the first proximal polymorphic variant with osteoarthritis is determined.

[0033] The methods described herein are useful for identifying or discovering additional polymorphic variants that may be used to further characterize a gene, region or loci associated with a condition, a disease (e.g., osteoarthritis), or a disorder. For example, allelotyping or genotyping data from the additional polymorphic variants may be used to identify a functional mutation or a region of linkage disequilibrium. In certain embodiments, polymorphic variants identified or discovered within a region comprising the first polymorphic variant associated with osteoarthritis are genotyped using the genetic methods and sample selection techniques described herein, and it can be determined whether those polymorphic variants are in linkage disequilibrium with the first polymorphic variant. The size of the region in linkage disequilibrium with the first polymorphic variant also can be assessed using these genotyping methods. Thus, provided herein are methods for determining whether a polymorphic variant is in linkage disequilibrium with a first polymorphic variant associated with osteoarthritis, and such information can be used in prognosis/diagnosis methods described herein.

Isolated Nucleic Acids

[0034] Featured herein are isolated *IL1RL1* nucleic acid variants depicted in SEQ ID NO: 1-4, and substantially identical nucleic acids thereof. A nucleic acid variant may be represented on one or both strands in a double-stranded nucleic acid or on one chromosomal complement (heterozygous) or both chromosomal complements (homozygous).

[0035] As used herein, the term "nucleic acid" includes DNA molecules (*e.g.*, a complementary DNA (cDNA) and genomic DNA (gDNA)) and RNA molecules (*e.g.*, mRNA, rRNA, siRNA and tRNA) and analogs of DNA or RNA, for example, by use of nucleotide analogs. The nucleic acid molecule can be single-stranded and it is often double-stranded. The term "isolated or purified nucleic acid" refers to nucleic acids that are separated from other nucleic acids present in the natural source of the nucleic acid. For example, with regard to genomic DNA, the term "isolated" includes nucleic acids which are separated from the chromosome with which the genomic DNA is naturally associated. An "isolated" nucleic acid is often free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and/or 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of 5' and/or 3' nucleotide sequences which flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. As used herein, the term "gene" refers to a nucleotide sequence that encodes a polypeptide.

[0036] Also included herein are nucleic acid fragments. These fragments often have a nucleotide sequence identical to a nucleotide sequence of SEQ ID NO: 1-4, a nucleotide sequence substantially identical to a nucleotide sequence of SEQ ID NO: 1-4, or a nucleotide sequence that is complementary to the foregoing. The nucleic acid fragment may be identical, substantially identical or homologous to a nucleotide sequence in an exon or an intron in a nucleotide sequence of SEQ ID NO: 1-4, and may encode a domain or part of a domain of a polypeptide. Sometimes, the fragment will comprises one or more of the polymorphic variations described herein as being associated with osteoarthritis. The nucleic acid fragment is often 50, 100, or 200 or fewer base pairs in length, and is sometimes about 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 2000, 3000, 4000, 5000, 10000, 15000, or 20000 base pairs in length. A nucleic acid fragment that is complementary to a nucleotide sequence identical or substantially identical to a nucleotide sequence in SEQ ID NO: 1-4 and hybridizes to such a nucleotide sequence under stringent conditions is often referred to as a "probe." Nucleic acid fragments often include one or more polymorphic sites, or sometimes have an end that is adjacent to a polymorphic site as described hereafter.

[0037] An example of a nucleic acid fragment is an oligonucleotide. As used herein, the term "oligonucleotide" refers to a nucleic acid comprising about 8 to about 50 covalently linked nucleotides, often comprising from about 8 to about 35 nucleotides, and more often from about 10 to about 25 nucleotides. The backbone and nucleotides within an oligonucleotide may be the same as those of naturally occurring nucleic acids, or analogs or derivatives of naturally occurring nucleic acids, provided that oligonucleotides having such analogs or derivatives retain the ability to hybridize specifically to a nucleic acid comprising a targeted polymorphism. Oligonucleotides described herein may be used as hybridization probes or as components of prognostic or diagnostic assays, for example, as described herein.

[0038] Oligonucleotides are typically synthesized using standard methods and equipment, such as the ABI™3900 High Throughput DNA Synthesizer and the EXPEDITE™ 8909 Nucleic Acid Synthesizer, both of which are available from Applied Biosystems (Foster City, CA). Analogs and derivatives are exemplified in U.S. Pat. Nos. 4,469,863; 5,536,821; 5,541,306; 5,637,683; 5,637,684; 5,700,922; 5,717,083; 5,719,262; 5,739,308; 5,773,601; 5,886,165; 5,929,226; 5,977,296; 6,140,482; WO 00/56746; WO 01/14398, and related publications. Methods for synthesizing oligonucleotides comprising such analogs or derivatives are disclosed, for example, in the patent publications cited above and in U.S. Pat. Nos. 5,614,622; 5,739,314; 5,955,599; 5,962,674; 6,117,992; in WO 00/75372; and in related publications.

[0039] Oligonucleotides may also be linked to a second moiety. The second moiety may be an additional nucleotide sequence such as a tail sequence (e.g., a polyadenosine tail), an adapter sequence (e.g., phage M13 universal tail sequence), and others. Alternatively, the second moiety may be a non-nucleotide moiety such as a moiety which facilitates linkage to a solid support or a label to facilitate detection of the oligonucleotide. Such labels include, without limitation, a radioactive label, a fluorescent label, a chemiluminescent label, a paramagnetic label, and the like. The second moiety may be attached to any position of the oligonucleotide, provided the oligonucleotide can hybridize to the nucleic acid comprising the polymorphism.

Uses for Nucleic Acid Sequence

[0040] Nucleic acid coding sequences may be used for diagnostic purposes for detection and control of polypeptide expression. Also, included herein are oligonucleotide sequences such as antisense RNA, small-interfering RNA (siRNA) and DNA molecules and ribozymes that function to inhibit translation of a polypeptide. Antisense techniques and RNA interference techniques are known in the art and are described herein.

[0041] Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme

9

molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, hammerhead motif ribozyme molecules may be engineered that specifically and efficiently catalyze endonucleolytic cleavage of RNA sequences corresponding to or complementary to *IL1RL1* nucleotide sequences. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between fifteen (15) and twenty (20) ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for predicted structural features such as secondary structure that may render the oligonucleotide sequence unsuitable. The suitability of candidate targets may also be evaluated by testing their accessibility to hybridization with complementary oligonucleotides, using ribonuclease protection assays.

[0042] Antisense RNA and DNA molecules, siRNA and ribozymes may be prepared by any method known in the art for the synthesis of RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides well known in the art such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

[0043] DNA encoding a polypeptide also may have a number of uses for the diagnosis of diseases, including osteoarthritis, resulting from aberrant expression of a target gene described herein. For example, the nucleic acid sequence may be used in hybridization assays of biopsies or autopsies to diagnose abnormalities of expression or function (*e.g.*, Southern or Northern blot analysis, in situ hybridization assays).

[0044] In addition, the expression of a polypeptide during embryonic development may also be determined using nucleic acid encoding the polypeptide. As addressed, *infra*, production of functionally impaired polypeptide is the cause of various disease states, such as osteoarthritis. *In situ* hybridizations using polypeptide as a probe may be employed to predict problems related to osteoarthritis. Further, as indicated, *infra*, administration of human active polypeptide, recombinantly produced as described herein, may be used to treat disease states related to functionally impaired polypeptide. Alternatively, gene therapy approaches may be employed to remedy deficiencies of functional polypeptide or to replace or compete with dysfunctional polypeptide.

Expression Vectors, Host Cells, and Genetically Engineered Cells

[0045] Provided herein are nucleic acid vectors, often expression vectors, which contain a *IL1RL1* nucleotide sequence, or a substantially identical sequence thereof. As used herein, the term "vector"

refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked and can include a plasmid, cosmid, or viral vector. The vector can be capable of autonomous replication or it can integrate into a host DNA. Viral vectors may include replication defective retroviruses, adenoviruses and adeno-associated viruses for example.

[0046] A vector can include a *IL1RL1* nucleotide sequence in a form suitable for expression of an encoded target polypeptide or target nucleic acid in a host cell. A "target polypeptide" is a polypeptide encoded by a *IL1RL1* nucleotide sequence, or a substantially identical nucleotide sequence thereof. The recombinant expression vector typically includes one or more regulatory sequences operatively linked to the nucleic acid sequence to be expressed. The term "regulatory sequence" includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence, as well as tissue-specific regulatory and/or inducible sequences. The design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of polypeptide desired, and the like. Expression vectors can be introduced into host cells to produce target polypeptides, including fusion polypeptides.

[0047] Recombinant expression vectors can be designed for expression of target polypeptides in prokaryotic or eukaryotic cells. For example, target polypeptides can be expressed in *E. coli*, insect cells (*e.g.*, using baculovirus expression vectors), yeast cells, or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

[0048] Expression of polypeptides in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion polypeptides. Fusion vectors add a number of amino acids to a polypeptide encoded therein, usually to the amino terminus of the recombinant polypeptide. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant polypeptide; 2) to increase the solubility of the recombinant polypeptide; and 3) to aid in the purification of the recombinant polypeptide by acting as a ligand in affinity purification. Often, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant polypeptide to enable separation of the recombinant polypeptide from the fusion moiety subsequent to purification of the fusion polypeptide. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith & Johnson, *Gene 67:* 31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding polypeptide, or polypeptide A, respectively, to the target recombinant polypeptide.

11

[0049] Purified fusion polypeptides can be used in screening assays and to generate antibodies specific for target polypeptides. In a therapeutic embodiment, fusion polypeptide expressed in a retroviral expression vector is used to infect bone marrow cells that are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (*e.g.*, six (6) weeks).

[0050] Expressing the polypeptide in host bacteria with an impaired capacity to proteolytically cleave the recombinant polypeptide is often used to maximize recombinant polypeptide expression (Gottesman, S., *Gene Expression Technology: Methods in Enzymology, Academic Press, San Diego, California 185:* 119-128 (1990)). Another strategy is to alter the nucleotide sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al., Nucleic Acids Res. 20:* 2111-2118 (1992)). Such alteration of nucleotide sequences can be carried out by standard DNA synthesis techniques.

[0051] When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. Recombinant mammalian expression vectors are often capable of directing expression of the nucleic acid in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Non-limiting examples of suitable tissue-specific promoters include an albumin promoter (liver-specific; Pinkert *et al., Genes Dev. 1:* 268-277 (1987)), lymphoid-specific promoters (Calame & Eaton, *Adv. Immunol. 43:* 235-275 (1988)), promoters of T cell receptors (Winoto & Baltimore, *EMBO J. 8:* 729-733 (1989)) promoters of immunoglobulins (Banerji *et al., Cell 33:* 729-740 (1983); Queen & Baltimore, *Cell 33:* 741-748 (1983)), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne & Ruddle, *Proc. Natl. Acad. Sci. USA 86:* 5473-5477 (1989)), pancreas-specific promoters (Edlund *et al., Science 230:* 912-916 (1985)), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are sometimes utilized, for example, the murine hox promoters (Kessel & Gruss, *Science 249:* 374-379 (1990)) and the α-fetopolypeptide promoter (Campes & Tilghman, *Genes Dev. 3:* 537-546 (1989)).

[0052] A *IL1RL1* nucleic acid also may be cloned into an expression vector in an antisense orientation. Regulatory sequences (*e.g.*, viral promoters and/or enhancers) operatively linked to a *IL1RL1* nucleic acid cloned in the antisense orientation can be chosen for directing constitutive, tissue specific or cell type specific expression of antisense RNA in a variety of cell types. Antisense expression vectors can be in the form of a recombinant plasmid, phagemid or attenuated virus. For a discussion of the regulation of gene expression using antisense genes *see, e.g.,* Weintraub *et al.,* Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics,* Vol. 1(1) (1986).

[0053] Also provided herein are host cells that include a *IL1RL1* nucleotide sequence within a recombinant expression vector or a fragment of such a nucleotide sequence which facilitate homologous recombination into a specific site of the host cell genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. Such terms refer not only to the particular subject cell but rather also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. A host cell can be any prokaryotic or eukaryotic cell. For example, a target polypeptide can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

[0054] Vectors can be introduced into host cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, transduction/infection, DEAE-dextran-mediated transfection, lipofection, or electroporation.

[0055] A host cell provided herein can be used to produce (*i.e.*, express) a target polypeptide or a substantially identical polypeptide thereof. Accordingly, further provided are methods for producing a target polypeptide using host cells described herein. In one embodiment, the method includes culturing host cells into which a recombinant expression vector encoding a target polypeptide has been introduced in a suitable medium such that a target polypeptide is produced. In another embodiment, the method further includes isolating a target polypeptide from the medium or the host cell.

[0056] Also provided are cells or purified preparations of cells which include a *IL1RL1* transgene, or which otherwise misexpress target polypeptide. Cell preparations can consist of human or non-human cells, *e.g.*, rodent cells, *e.g.*, mouse or rat cells, rabbit cells, or pig cells. In preferred embodiments, the cell or cells include a *IL1RL1* transgene (*e.g.*, a heterologous form of a *IL1RL1* gene, such as a human gene expressed in non-human cells). The transgene can be misexpressed, *e.g.*, overexpressed or underexpressed. In other preferred embodiments, the cell or cells include a gene which misexpress an endogenous target polypeptide (*e.g.*, expression of a gene is disrupted, also known as a knockout). Such cells can serve as a model for studying disorders which are related to mutated or mis-expressed alleles or for use in drug screening. Also provided are human cells (*e.g.*, a hematopoietic stem cells) transfected with a *IL1RL1* nucleic acid.

[0057] Also provided are cells or a purified preparation thereof (*e.g.*, human cells) in which an endogenous *IL1RL1* nucleic acid is under the control of a regulatory sequence that does not normally control the expression of the endogenous gene. The expression characteristics of an endogenous gene within a cell (*e.g.*, a cell line or microorganism) can be modified by inserting a heterologous DNA

13

regulatory element into the genome of the cell such that the inserted regulatory element is operably linked to the corresponding endogenous gene. For example, an endogenous corresponding gene (e.g., a gene which is "transcriptionally silent," not normally expressed, or expressed only at very low levels) may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell. Techniques such as targeted homologous recombinations, can be used to insert the heterologous DNA as described in, e.g., Chappel, US 5,272,071; WO 91/06667, published on May 16, 1991.

Transgenic Animals

[0058] Non-human transgenic animals that express a heterologous target polypeptide (e.g., expressed from a *IL1RL1* nucleic acid or substantially identical sequence thereof) can be generated. Such animals are useful for studying the function and/or activity of a target polypeptide and for identifying and/or evaluating modulators of the activity of *IL1RL1* nucleic acids and encoded polypeptides. As used herein, a "transgenic animal" is a non-human animal such as a mammal (e.g., a non-human primate such as chimpanzee, baboon, or macaque; an ungulate such as an equine, bovine, or caprine; or a rodent such as a rat, a mouse, or an Israeli sand rat), a bird (e.g., a chicken or a turkey), an amphibian (e.g., a frog, salamander, or newt), or an insect (e.g., *Drosophila melanogaster*), in which one or more of the cells of the animal includes a transgene. A transgene is exogenous DNA or a rearrangement (e.g., a deletion of endogenous chromosomal DNA) that is often integrated into or occurs in the genome of cells in a transgenic animal. A transgene can direct expression of an encoded gene product in one or more cell types or tissues of the transgenic animal, and other transgenes can reduce expression (e.g., a knockout). Thus, a transgenic animal can be one in which an endogenous nucleic acid homologous to a *IL1RL1* nucleic acid has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal (e.g., an embryonic cell of the animal) prior to development of the animal.

[0059] Intronic sequences and polyadenylation signals can also be included in the transgene to increase expression efficiency of the transgene. One or more tissue-specific regulatory sequences can be operably linked to a *IL1RL1* nucleotide sequence to direct expression of an encoded polypeptide to particular cells. A transgenic founder animal can be identified based upon the presence of a *IL1RL1* nucleotide sequence in its genome and/or expression of encoded mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a *IL1RL1* nucleotide sequence can further be bred to other transgenic animals carrying other transgenes.

[0060] Target polypeptides can be expressed in transgenic animals or plants by introducing, for example, a *IL1RL1* nucleic acid into the genome of an animal that encodes the target polypeptide. In

14

preferred embodiments the nucleic acid is placed under the control of a tissue specific promoter, *e.g.*, a milk or egg specific promoter, and recovered from the milk or eggs produced by the animal. Also included is a population of cells from a transgenic animal.

Target Polypeptides

[0061] Also featured herein are isolated target polypeptides, which are encoded by a *IL1RL1* nucleotide sequence (e.g., SEQ ID NO: 1-4), or a substantially identical nucleotide sequence thereof. Examples of *IL1RL1* polypeptides are set forth in SEQ ID NO: 5-7. The term "polypeptide" as used herein includes proteins and peptides. An "isolated" or "purified" polypeptide or protein is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. In one embodiment, the language "substantially free" means preparation of a target polypeptide having less than about 30%, 20%, 10% and more preferably 5% (by dry weight), of non-target polypeptide (also referred to herein as a "contaminating protein"), or of chemical precursors or non-target chemicals. When the target polypeptide or a biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, specifically, where culture medium represents less than about 20%, sometimes less than about 10%, and often less than about 5% of the volume of the polypeptide preparation. Isolated or purified target polypeptide preparations are sometimes 0.01 milligrams or more or 0.1 milligrams or more, and often 1.0 milligrams or more and 10 milligrams or more in dry weight.

[0062] Further included herein are target polypeptide fragments. The polypeptide fragment may be a domain or part of a domain of a target polypeptide. The polypeptide fragment may have increased, decreased or unexpected biological activity. The polypeptide fragment is often 50 or fewer, 100 or fewer, or 200 or fewer amino acids in length, and is sometimes 300, 400, 500, 600, 700, or 900 or fewer amino acids in length.

[0063] Interleukin 1 receptor-like 1 isoform 1 (SEQ ID NO: 5) is a member of the interleukin 1 receptor family with no known ligand (orphan receptor). *IL1RL1* exists in soluble (SEQ ID NO: 6-7) and transmembrane forms, suggesting that it may have ligand or ligand scavenging activity. In an embodiment, *IL1RL1* protein agents may be administered to treat or prevent the occurrence of OA. *IL1RL1* protein agents include *IL1RL1* polypeptides or fragments thereof that have *IL1RL1* ligand activity (e.g., recombinant polypeptides of SEQ ID NO: 6-7). In a related embodiment, *IL1RL1* protein agents include *IL1RL1* polypeptides or fragments thereof that have *IL1RL1* ligand scavenging activity (e.g., recombinant polypeptide of SEQ ID NO: 5). Isolated *IL1RL1* polypeptides featured herein include the full-length polypeptide, the mature polypeptide (i.e., the polypeptide without the signal sequence

15

MGFWILAILTILMYSTAA) or a polypeptide fragment containing a domain or part of a *IL1RL1* domain. The polypeptide fragment may have increased, decreased or unexpected biological activity.

[0064] Substantially identical target polypeptides may depart from the amino acid sequences of target polypeptides in different manners. For example, conservative amino acid modifications may be introduced at one or more positions in the amino acid sequences of target polypeptides. A "conservative amino acid substitution" is one in which the amino acid is replaced by another amino acid having a similar structure and/or chemical function. Families of amino acid residues having similar structures and functions are well known. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Also, essential and non-essential amino acids may be replaced. A "non-essential" amino acid is one that can be altered without abolishing or substantially altering the biological function of a target polypeptide, whereas altering an "essential" amino acid abolishes or substantially alters the biological function of a target polypeptide. Amino acids that are conserved among target polypeptides are typically essential amino acids. In certain embodiments, the polypeptide includes one or more non-synonymous polymorphic variants associated with osteoarthritis.

[0065] Also, target polypeptides may exist as chimeric or fusion polypeptides. As used herein, a target "chimeric polypeptide" or target "fusion polypeptide" includes a target polypeptide linked to a non-target polypeptide. A "non-target polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a polypeptide which is not substantially identical to the target polypeptide, which includes, for example, a polypeptide that is different from the target polypeptide and derived from the same or a different organism. The target polypeptide in the fusion polypeptide can correspond to an entire or nearly entire target polypeptide or a fragment thereof. The non-target polypeptide can be fused to the N-terminus or C-terminus of the target polypeptide.

[0066] Fusion polypeptides can include a moiety having high affinity for a ligand. For example, the fusion polypeptide can be a GST-target fusion polypeptide in which the target sequences are fused to the C-terminus of the GST sequences, or a polyhistidine-target fusion polypeptide in which the target polypeptide is fused at the N- or C-terminus to a string of histidine residues. Such fusion polypeptides can facilitate purification of recombinant target polypeptide. Expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide), and a nucleotide sequence in SEQ ID NO: 1-4, or a substantially identical nucleotide sequence thereof, can be cloned into an expression vector such that the fusion moiety is linked in-frame to the target polypeptide. Further, the fusion polypeptide can be a target polypeptide containing a heterologous signal sequence at its N-

16

terminus. In certain host cells (*e.g.*, mammalian host cells), expression, secretion, cellular internalization, and cellular localization of a target polypeptide can be increased through use of a heterologous signal sequence. Fusion polypeptides can also include all or a part of a serum polypeptide (*e.g.*, an IgG constant region or human serum albumin).

[0067] Target polypeptides can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. Administration of these target polypeptides can be used to affect the bioavailability of a substrate of the target polypeptide and may effectively increase target polypeptide biological activity in a cell. Target fusion polypeptides may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a target polypeptide; (ii) mis-regulation of the gene encoding the target polypeptide; and (iii) aberrant post-translational modification of a target polypeptide. Also, target polypeptides can be used as immunogens to produce anti-target antibodies in a subject, to purify target polypeptide ligands or binding partners, and in screening assays to identify molecules which inhibit or enhance the interaction of a target polypeptide with a substrate.

[0068] In addition, polypeptides can be chemically synthesized using techniques known in the art (See, *e.g.*, Creighton, 1983 Proteins. New York, N.Y.: W. H. Freeman and Company; and Hunkapiller et al., (1984) Nature July 12 -18;310(5973):105-11). For example, a relative short fragment can be synthesized by use of a peptide synthesizer. Furthermore, if desired, non-classical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the fragment sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid, a-amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, g-Abu, e-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, b-alanine, fluoroamino acids, designer amino acids such as b-methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

[0069] Polypeptides and polypeptide fragments sometimes are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH4; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; and the like. Additional post-translational modifications include, for example, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell

expression. The polypeptide fragments may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the polypeptide.

[0070] Also provided are chemically modified derivatives of polypeptides that can provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see e.g., U.S. Pat. No: 4,179,337. The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

[0071] The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

[0072] The polymers should be attached to the polypeptide with consideration of effects on functional or antigenic domains of the polypeptide. There are a number of attachment methods available to those skilled in the art (e.g., EP 0 401 384 (coupling PEG to G-CSF) and Malik et al. (1992) Exp Hematol. September;20(8):1028-35 (pegylation of GM-CSF using tresyl chloride)). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues, glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. For therapeutic purposes, the attachment sometimes is at an amino group, such as attachment at the N-terminus or lysine group.

[0073] Proteins can be chemically modified at the N-terminus. Using polyethylene glycol as an illustration of such a composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, and the like), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus may be accomplished by

reductive alkylation, which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

Substantially Identical Nucleic Acids and Polypeptides

[0074] Nucleotide sequences and polypeptide sequences that are substantially identical to a *IL1RL1* nucleotide sequence and the target polypeptide sequences encoded by those nucleotide sequences, respectively, are included herein. The term "substantially identical" as used herein refers to two or more nucleic acids or polypeptides sharing one or more identical nucleotide sequences or polypeptide sequences, respectively. Included are nucleotide sequences or polypeptide sequences that are 55% or more, 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 95% or more (each often within a 1%, 2%, 3% or 4% variability) identical to a *IL1RL1* nucleotide sequence or the encoded target polypeptide amino acid sequences. One test for determining whether two nucleic acids are substantially identical is to determine the percent of identical nucleotide sequences or polypeptide sequences shared between the nucleic acids or polypeptides.

[0075] Calculations of sequence identity are often performed as follows. Sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). The length of a reference sequence aligned for comparison purposes is sometimes 30% or more, 40% or more, 50% or more, often 60% or more, and more often 70% or more, 80% or more, 90% or more, or 100% of the length of the reference sequence. The nucleotides or amino acids at corresponding nucleotide or polypeptide positions, respectively, are then compared among the two sequences. When a position in the first sequence is occupied by the same nucleotide or amino acid as the corresponding position in the second sequence, the nucleotides or amino acids are deemed to be identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, introduced for optimal alignment of the two sequences.

[0076] Comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. Percent identity between two amino acid or nucleotide sequences can be determined using the algorithm of Meyers & Miller, *CABIOS 4:* 11-17 (1989), which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4. Also, percent identity between two amino acid sequences can be determined using the Needleman & Wunsch, *J. Mol. Biol. 48:* 444-453 (1970) algorithm which has been incorporated into the GAP program in the GCG software package (available at

19

the http address www.gcg.com), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. Percent identity between two nuclcotide scquences can be determined using the GAP program in the GCG software package (available at http address www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A set of parameters often used is a Blossum 62 scoring matrix with a gap open penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

[0077] Another manner for determining if two nucleic acids are substantially identical is to assess whether a polynucleotide homologous to one nucleic acid will hybridize to the other nucleic acid under stringent conditions. As use herein, the term "stringent conditions" refers to conditions for hybridization and washing. Stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. , 6.3.1-6.3.6 (1989). Aqueous and non-aqueous methods are described in that reference and either can be used. An example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50°C. Another example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 55°C. A further example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 60°C. Often, stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 65°C. More often, stringency conditions are 0.5M sodium phosphate, 7% SDS at 65°C, followed by one or more washes at 0.2X SSC, 1% SDS at 65°C.

[0078] An example of a substantially identical nucleotide sequence to a nucleotide sequence in SEQ ID NO: 1-4 is one that has a different nucleotide sequence but still encodes the same polypeptide sequence encoded by the nucleotide sequence in SEQ ID NO: 1-4. Another example is a nucleotide sequence that encodes a polypeptide having a polypeptide sequence that is more than 70% or more identical to, sometimes more than 75% or more, 80% or more, or 85% or more identical to, and often more than 90% or more and 95% or more identical to a polypeptide sequence encoded by a nucleotide sequence in SEQ ID NO: 1-4.

[0079] Nucleotide sequences in SEQ ID NO: 1-4 and amino acid sequences of encoded polypeptides can be used as "query sequences" to perform a search against public databases to identify other family members or related sequences, for example. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul *et al., J. Mol. Biol. 215:* 403-10 (1990). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to nucleotide sequences in SEQ ID NO: 1-4. BLAST polypeptide

20

searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to polypeptides encoded by the nucleotide sequences of SEQ ID NO: 1-4. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al., Nucleic Acids Res. 25(17):* 3389-3402 (1997). When utilizing BLAST and Gapped BLAST programs, default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used (*see* the http address www.ncbi.nlm.nih.gov).

[0080] A nucleic acid that is substantially identical to a nucleotide sequence in SEQ ID NO: 1-4 may include polymorphic sites at positions equivalent to those described herein when the sequences are aligned. For example, using the alignment procedures described herein, SNPs in a sequence substantially identical to a sequence in SEQ ID NO: 1-4 can be identified at nucleotide positions that match (*i.e.*, align) with nucleotides at SNP positions in each nucleotide sequence in SEQ ID NO: 1-4. Also, where a polymorphic variation results in an insertion or deletion, insertion or deletion of a nucleotide sequence from a reference sequence can change the relative positions of other polymorphic sites in the nucleotide sequence.

[0081] Substantially identical nucleotide and polypeptide sequences include those that are naturally occurring, such as allelic variants (same locus), splice variants, homologs (different locus), and orthologs (different organism) or can be non-naturally occurring. Non-naturally occurring variants can be generated by mutagenesis techniques, including those applied to polynucleotides, cells, or organisms. The variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared in the encoded product). Orthologs, homologs, allelic variants, and splice variants can be identified using methods known in the art. These variants normally comprise a nucleotide sequence encoding a polypeptide that is 50% or more, about 55% or more, often about 70-75% or more or about 80-85% or more, and sometimes about 90-95% or more identical to the amino acid sequences of target polypeptides or a fragment thereof. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions to a nucleotide sequence in SEQ ID NO: 1-4 or a fragment of this sequence. Nucleic acid molecules corresponding to orthologs, homologs, and allelic variants of a nucleotide sequence in SEQ ID NO: 1-4 can further be identified by mapping the sequence to the same chromosome or locus as the nucleotide sequence in SEQ ID NO: 1-4.

[0082] Also, substantially identical nucleotide sequences may include codons that are altered with respect to the naturally occurring sequence for enhancing expression of a target polypeptide in a particular expression system. For example, the nucleic acid can be one in which one or more codons are altered, and often 10% or more or 20% or more of the codons are altered for optimized expression in

bacteria (*e.g.*, *E. coli.*), yeast (*e.g.*, *S. cervesiae*), human (*e.g.*, 293 cells), insect, or rodent (*e.g.*, hamster) cells.

Methods for Identifying Risk of Osteoarthritis

[0083] Methods for prognosing and diagnosing osteoarthritis are included herein. These methods include detecting the presence or absence of one or more polymorphic variations in a nucleotide sequence associated with osteoarthritis, such as variants in or around the loci set forth herein, or a substantially identical sequence thereof, in a sample from a subject, where the presence of a polymorphic variant described herein is indicative of a risk of osteoarthritis. Determining a risk of osteoarthritis sometimes refers to determining whether an individual is at an increased risk of osteoarthritis (*e.g.*, intermediate risk or higher risk).

[0084] Thus, featured herein is a method for identifying a subject who is at risk of osteoarthritis, which comprises detecting an aberration associated with osteoarthritis in a nucleic acid sample from the subject. An embodiment is a method for detecting a risk of osteoarthritis in a subject, which comprises detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject, where the nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) a nucleotide sequence of SEQ ID NO: 1-4; (b) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-4; (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-4, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-4; and (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising the polymorphic site; whereby the presence of the polymorphic variation is indicative of a predisposition to osteoarthritis in the subject. In certain embodiments, polymorphic variants at the positions described herein are detected for determining a risk of osteoarthritis, and polymorphic variants at positions in linkage disequilibrium with these positions are detected for determining a risk of osteoarthritis. As used herein, "SEQ ID NO: 1-4" refers to individual sequences in SEQ ID NO: 1, 2, 3 or 4, each sequence being separately applicable to embodiments described herein.

[0085] Risk of osteoarthritis sometimes is expressed as a probability, such as an odds ratio, percentage, or risk factor. Risk often is based upon the presence or absence of one or more polymorphic variants described herein, and also may be based in part upon phenotypic traits of the individual being tested. Methods for calculating risk based upon patient data are well known (*see, e.g.*, Agresti, *Categorical Data Analysis*, 2nd Ed. 2002. Wiley). Allelotyping and genotyping analyses may be carried out in populations other than those exemplified herein to enhance the predictive power of the prognostic

method. These further analyses are executed in view of the exemplified procedures described herein, and may be based upon the same polymorphic variations or additional polymorphic variations.

[0086] In certain embodiments, determining the presence of a combination of two or more polymorphic variants associated with osteoarthritis in one or more genetic loci (e.g., one or more genes) of the sample is determined to identify, quantify and/or estimate, risk of osteoarthritis. The risk often is the probability of having or developing osteoarthritis. The risk sometimes is expressed as a relative risk with respect to a population average risk of osteoarthritis, and sometimes is expressed as a relative risk with respect to the lowest risk group. Such relative risk assessments often are based upon penetrance values determined by statistical methods, and are particularly useful to clinicians and insurance companies for assessing risk of osteoarthritis (e.g., a clinician can target appropriate detection, prevention and therapeutic regimens to a patient after determining the patient's risk of osteoarthritis, and an insurance company can fine tune actuarial tables based upon population genotype assessments of osteoarthritis risk). Risk of osteoarthritis sometimes is expressed as an odds ratio, which is the odds of a particular person having a genotype has or will develop osteoarthritis with respect to another genotype group (e.g., the most disease protective genotype or population average). In related embodiments, the determination is utilized to identify a subject at risk of osteoarthritis. In an embodiment, two or more polymorphic variations are detected in two or more regions in human genomic DNA associated with increased risk of osteoarthritis, such as a locus containing a *IL1RL1*, for example. In certain embodiments, 3 or more, or 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more polymorphic variants are detected in the sample. In specific embodiments, polymorphic variants are detected in a *IL1RL1* region, for example. In certain embodiments, polymorphic variants are detected at other genetic loci (e.g., the polymorphic variants can be detected in *IL1RL1* in addition to other loci or only in other loci), where the other loci include but are not limited to those described in concurrently-filed patent applications having attorney docket number 524593008700, 524593008800, 524593008900, 524593009000 or 524593009200, which is incorporated herein by reference in its entirety.

[0087] Results from prognostic tests may be combined with other test results to diagnose osteoarthritis. For example, prognostic results may be gathered, a patient sample may be ordered based on a determined predisposition to osteoarthritis, the patient sample is analyzed, and the results of the analysis may be utilized to diagnose osteoarthritis. Also osteoarthritis diagnostic method can be developed from studies used to generate prognostic methods in which populations are stratified into subpopulations having different progressions of osteoarthritis. In another embodiment, prognostic results may be gathered, a patient's risk factors for developing osteoarthritis (*e.g.*, age, weight, race, diet) analyzed, and a patient sample may be ordered based on a determined predisposition to osteoarthritis.

[0088] The nucleic acid sample typically is isolated from a biological sample obtained from a subject. For example, nucleic acid can be isolated from blood, saliva, sputum, urine, cell scrapings, and

biopsy tissue. The nucleic acid sample can be isolated from a biological sample using standard techniques, such as the technique described in Example 2. As used herein, the term "subject" refers primarily to humans but also refers to other mammals such as dogs, cats, and ungulates (e.g., cattle, sheep, and swine). Subjects also include avians (e.g., chickens and turkeys), reptiles, and fish (e.g., salmon), as embodiments described herein can be adapted to nucleic acid samples isolated from any of these organisms. The nucleic acid sample may be isolated from the subject and then directly utilized in a method for determining the presence of a polymorphic variant, or alternatively, the sample may be isolated and then stored (e.g., frozen) for a period of time before being subjected to analysis.

[0089] The presence or absence of a polymorphic variant is determined using one or both chromosomal complements represented in the nucleic acid sample. Determining the presence or absence of a polymorphic variant in both chromosomal complements represented in a nucleic acid sample from a subject having a copy of each chromosome is useful for determining the zygosity of an individual for the polymorphic variant (i.e., whether the individual is homozygous or heterozygous for the polymorphic variant). Any oligonucleotide-based diagnostic may be utilized to determine whether a sample includes the presence or absence of a polymorphic variant in a sample. For example, primer extension methods, ligase sequence determination methods (e.g., U.S. Pat. Nos. 5,679,524 and 5,952,174, and WO 01/27326), mismatch sequence determination methods (e.g., U.S. Pat. Nos. 5,851,770; 5,958,692; . 6,110,684; and 6,183,958), microarray sequence determination methods, restriction fragment length polymorphism (RFLP), single strand conformation polymorphism detection (SSCP) (e.g., U.S. Pat. Nos. 5,891,625 and 6,013,499), PCR-based assays (e.g., TAQMAN® PCR System (Applied Biosystems)), and nucleotide sequencing methods may be used.

[0090] Oligonucleotide extension methods typically involve providing a pair of oligonucleotide primers in a polymerase chain reaction (PCR) or in other nucleic acid amplification methods for the purpose of amplifying a region from the nucleic acid sample that comprises the polymorphic variation. One oligonucleotide primer is complementary to a region 3' of the polymorphism and the other is complementary to a region 5' of the polymorphism. A PCR primer pair may be used in methods disclosed in U.S. Pat. Nos. 4,683,195; 4,683,202, 4,965,188; 5,656,493; 5,998,143; 6,140,054; WO 01/27327; and WO 01/27329 for example. PCR primer pairs may also be used in any commercially available machines that perform PCR, such as any of the GENEAMP® Systems available from Applied Biosystems. Also, those of ordinary skill in the art will be able to design oligonucleotide primers based upon a IL1RL1 nucleotide sequence using knowledge available in the art.

[0091] Also provided is an extension oligonucleotide that hybridizes to the amplified fragment adjacent to the polymorphic variation. As used herein, the term "adjacent" refers to the 3' end of the extension oligonucleotide being often 1 nucleotide from the 5' end of the polymorphic site, and sometimes 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from the 5' end of the polymorphic site, in the nucleic

acid when the extension oligonucleotide is hybridized to the nucleic acid. The extension oligonucleotide then is extended by one or more nucleotides, and the number and/or type of nucleotides that are added to the extension oligonucleotide determine whether the polymorphic variant is present. Oligonucleotide extension methods are disclosed, for example, in U.S. Pat. Nos. 4,656,127; 4,851,331; 5,679,524; 5,834,189; 5,876,934; 5,908,755; 5,912,118; 5,976,802; 5,981,186; 6,004,744; 6,013,431; 6,017,702; 6,046,005; 6,087,095; 6,210,891; and WO 01/20039. Oligonucleotide extension methods using mass spectrometry are described, for example, in U.S. Pat. Nos. 5,547,835; 5,605,798; 5,691,141; 5,849,542; 5,869,242; 5,928,906; 6,043,031; and 6,194,144, and a method often utilized is described herein in Example 2.

[0092] A microarray can be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A microarray may include any oligonucleotides described herein, and methods for making and using oligonucleotide microarrays suitable for diagnostic use are disclosed in U.S. Pat. Nos. 5,492,806; 5,525,464; 5,589,330; 5,695,940; 5,849,483; 6,018,041; 6,045,996; 6,136,541; 6,142,681; 6,156,501; 6,197,506; 6,223,127; 6,225,625; 6,229,911; 6,239,273; WO 00/52625; WO 01/25485; and WO 01/29259. The microarray typically comprises a solid support and the oligonucleotides may be linked to this solid support by covalent bonds or by non-covalent interactions. The oligonucleotides may also be linked to the solid support directly or by a spacer molecule. A microarray may comprise one or more oligonucleotides complementary to a polymorphic site set forth herein.

[0093] A kit also may be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A kit often comprises one or more pairs of oligonucleotide primers useful for amplifying a fragment of a nucleotide sequence of SEQ ID NO: 1-4 or a substantially identical sequence thereof, where the fragment includes a polymorphic site. The kit sometimes comprises a polymerizing agent, for example, a thermostable nucleic acid polymerase such as one disclosed in U.S. Pat. Nos. 4,889,818 or 6,077,664. Also, the kit often comprises an elongation oligonucleotide that hybridizes to a *IL1RL1* nucleotide sequence in a nucleic acid sample adjacent to the polymorphic site. Where the kit includes an elongation oligonucleotide, it also often comprises chain elongating nucleotides, such as dATP, dTTP, dGTP, dCTP, and dITP, including analogs of dATP, dTTP, dGTP, dCTP and dITP, provided that such analogs are substrates for a thermostable nucleic acid polymerase and can be incorporated into a nucleic acid chain elongated from the extension oligonucleotide. Along with chain elongating nucleotides would be one or more chain terminating nucleotides such as ddATP, ddTTP, ddGTP, ddCTP, and the like. In an embodiment, the kit comprises one or more oligonucleotide primer pairs, a polymerizing agent, chain elongating nucleotides, at least one elongation oligonucleotide, and one or more chain terminating nucleotides. Kits optionally include buffers, vials, microtiter plates, and instructions for use.

25

[0094] An individual identified as being at risk of osteoarthritis may be heterozygous or homozygous with respect to the allele associated with a higher risk of osteoarthritis. A subject homozygous for an allele associated with an increased risk of osteoarthritis is at a comparatively high risk of osteoarthritis, a subject heterozygous for an allele associated with an increased risk of osteoarthritis is at a comparatively intermediate risk of osteoarthritis, and a subject homozygous for an allele associated with a decreased risk of osteoarthritis is at a comparatively low risk of osteoarthritis. A genotype may be assessed for a complementary strand, such that the complementary nucleotide at a particular position is detected.

[0095] Also featured are methods for determining risk of osteoarthritis and/or identifying a subject at risk of osteoarthritis by contacting a polypeptide or protein encoded by a *IL1RL1* nucleotide sequence from a subject with an antibody that specifically binds to an epitope associated with increased risk of osteoarthritis in the polypeptide (e.g., an epitope comprising an alanine at position 78 in an *IR1LR1* polypeptide).

Applications of Prognostic and Diagnostic Results to Pharmacogenomic Methods

[0096] Pharmacogenomics is a discipline that involves tailoring a treatment for a subject according to the subject's genotype as a particular treatment regimen may exert a differential effect depending upon the subject's genotype. For example, based upon the outcome of a prognostic test described herein, a clinician or physician may target pertinent information and preventative or therapeutic treatments to a subject who would be benefited by the information or treatment and avoid directing such information and treatments to a subject who would not be benefited (*e.g.*, the treatment has no therapeutic effect and/or the subject experiences adverse side effects).

[0097] The following is an example of a pharmacogenomic embodiment. A particular treatment regimen can exert a differential effect depending upon the subject's genotype. Where a candidate therapeutic exhibits a significant interaction with a major allele and a comparatively weak interaction with a minor allele (*e.g.*, an order of magnitude or greater difference in the interaction), such a therapeutic typically would not be administered to a subject genotyped as being homozygous for the minor allele, and sometimes not administered to a subject genotyped as being heterozygous for the minor allele. In another example, where a candidate therapeutic is not significantly toxic when administered to subjects who are homozygous for a major allele but is comparatively toxic when administered to subjects heterozygous or homozygous for a minor allele, the candidate therapeutic is not typically administered to subjects who are genotyped as being heterozygous or homozygous with respect to the minor allele.

[0098] The methods described herein are applicable to pharmacogenomic methods for preventing, alleviating or treating osteoarthritis. For example, a nucleic acid sample from an individual may be subjected to a prognostic test described herein. Where one or more polymorphic variations associated

26

with increased risk of osteoarthritis are identified in a subject, information for preventing or treating osteoarthritis and/or one or more osteoarthritis treatment regimens then may be prescribed to that subject.

[0099] In certain embodiments, a treatment or preventative regimen is specifically prescribed and/or administered to individuals who will most benefit from it based upon their risk of developing osteoarthritis assessed by the methods described herein. Thus, provided are methods for identifying a subject predisposed to osteoarthritis and then prescribing a therapeutic or preventative regimen to individuals identified as having a predisposition. Thus, certain embodiments are directed to a method for reducing osteoarthritis in a subject, which comprises: detecting the presence or absence of a polymorphic variant associated with osteoarthritis in a nucleotide sequence in a nucleic acid sample from a subject, where the nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) a nucleotide sequence of SEQ ID NO: 1-4; (b) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-4; (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-4, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-4; and (d) a fragment of a polynucleotide sequence of (a), (b), or (c); and prescribing or administering a treatment regimen to a subject from whom the sample originated where the presence of a polymorphic variation associated with osteoarthritis is detected in the nucleotide sequence. In these methods, predisposition results may be utilized in combination with other test results to diagnose osteoarthritis.

[0100] Certain preventative treatments often are prescribed to subjects having a predisposition to osteoarthritis and where the subject is diagnosed with osteoarthritis or is diagnosed as having symptoms indicative of an early stage of osteoarthritis. The treatment sometimes is preventative (e.g., is prescribed or administered to reduce the probability that osteoarthritis arises or progresses), sometimes is therapeutic, and sometimes delays, alleviates or halts the progression of osteoarthritis. Any known preventative or therapeutic treatment for alleviating or preventing the occurrence of osteoarthritis is prescribed and/or administered. For example, the treatment often is directed to decreasing pain and improving joint movement. Examples of OA treatments include exercises to keep joints flexible and improve muscle strength. Different medications to control pain, including corticosteroids and nonsteroidal anti-inflammatory drugs (NSAIDs, e.g., Voltaren); cyclooxygenase-2 (COX-2) inhibitors (e.g., Celebrex, Vioxx, Mobic, and Bextra); monoclonal antibodies (e.g., Remicade); tumor necrosis factor inhibitors (e.g., Enbrel); or injections of glucocorticoids, hyaluronic acid or chondrotin sulfate into joints that are inflamed and not responsive to NSAIDS. Orally administered chondroitin sulfate also may be used as a therapeutic, as it may increase hyaluronic acid levels and viscosity of synovial fluid, and decrease collagenase levels in synovial fluid. Also, glucosamine can serve as an OA therapeutic as delivering it into joints may inhibit enzymes involved in cartilage degradation and enhance the

27

production of hyaluronic acid. For mild pain without inflammation, acetaminophen may be used. Other treatments include: heat/cold therapy for temporary pain relief; joint protection to prevent strain or stress on painful joints; surgery to relieve chronic pain in damaged joints; and weight control to prevent extra stress on weight-bearing joints.

[0101] As therapeutic approaches for treating osteoarthritis continue to evolve and improve, the goal of treatments for osteoarthritis related disorders is to intervene even before clinical signs first manifest. Thus, genetic markers associated with susceptibility to osteoarthritis prove useful for early diagnosis, prevention and treatment of osteoarthritis.

[0102] As osteoarthritis preventative and treatment information can be specifically targeted to subjects in need thereof (e.g., those at risk of developing osteoarthritis or those in an early stage of osteoarthritis), provided herein is a method for preventing or reducing the risk of developing osteoarthritis in a subject, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying a subject with a predisposition to osteoarthritis, whereby the presence of the polymorphic variation is indicative of a predisposition to osteoarthritis in the subject; and (c) if such a predisposition is identified, providing the subject with information about methods or products to prevent or reduce osteoarthritis or to delay the onset of osteoarthritis. Also provided is a method of targeting information or advertising to a subpopulation of a human population based on the subpopulation being genetically predisposed to a disease or condition, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying the subpopulation of subjects in which the polymorphic variation is associated with osteoarthritis; and (c) providing information only to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition.

[0103] Pharmacogenomics methods also may be used to analyze and predict a response to osteoarthritis treatment or a drug. For example, if pharmacogenomics analysis indicates a likelihood that an individual will respond positively to osteoarthritis treatment with a particular drug, the drug may be administered to the individual. Conversely, if the analysis indicates that an individual is likely to respond negatively to treatment with a particular drug, an alternative course of treatment may be prescribed. A negative response may be defined as either the absence of an efficacious response or the presence of toxic side effects. The response to a therapeutic treatment can be predicted in a background study in which subjects in any of the following populations are genotyped: a population that responds favorably to a treatment regimen, a population that does not respond significantly to a treatment regimen, and a population that responds adversely to a treatment regimen (e.g., exhibits one or more side effects). These populations are provided as examples and other populations and subpopulations may be analyzed. Based

28

upon the results of these analyses, a subject is genotyped to predict whether he or she will respond favorably to a treatment regimen, not respond significantly to a treatment regimen, or respond adversely to a treatment regimen.

[0104] The tests described herein also are applicable to clinical drug trials. One or more polymorphic variants indicative of response to an agent for treating osteoarthritis or to side effects to an agent for treating osteoarthritis may be identified using the methods described herein. Thereafter, potential participants in clinical trials of such an agent may be screened to identify those individuals most likely to respond favorably to the drug and exclude those likely to experience side effects. In that way, the effectiveness of drug treatment may be measured in individuals who respond positively to the drug, without lowering the measurement as a result of the inclusion of individuals who are unlikely to respond positively in the study and without risking undesirable safety problems.

[0105] Thus, another embodiment is a method of selecting an individual for inclusion in a clinical trial of a treatment or drug comprising the steps of: (a) obtaining a nucleic acid sample from an individual; (b) determining the identity of a polymorphic variation which is associated with a positive response to the treatment or the drug, or at least one polymorphic variation which is associated with a negative response to the treatment or the drug in the nucleic acid sample, and (c) including the individual in the clinical trial if the nucleic acid sample contains said polymorphic variation associated with a positive response to the treatment or the drug or if the nucleic acid sample lacks said polymorphic variation associated with a negative response to the treatment or the drug. In addition, the methods described herein for selecting an individual for inclusion in a clinical trial of a treatment or drug encompass methods with any further limitation described in this disclosure, or those following, specified alone or in any combination. The polymorphic variation may be in a sequence selected individually or in any combination from the group consisting of (i) a nucleotide sequence of SEQ ID NO: 1-4; (ii) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-4; (iii) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-4, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-4; and (iv) a fragment of a polynucleotide sequence of (i), (ii), or (iii) comprising the polymorphic site. The including step (c) optionally comprises administering the drug or the treatment to the individual if the nucleic acid sample contains the polymorphic variation associated with a positive response to the treatment or the drug and the nucleic acid sample lacks said biallelic marker associated with a negative response to the treatment or the drug.

[0106] Also provided herein is a method of partnering between a diagnostic/prognostic testing provider and a provider of a consumable product, which comprises: (a) the diagnostic/prognostic testing provider detects the presence or absence of a polymorphic variation associated with osteoarthritis at a

polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) the diagnostic/prognostic testing provider identifies the subpopulation of subjects in which the polymorphic variation is associated with osteoarthritis; (c) the diagnostic/prognostic testing provider forwards information to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition; and (d) the provider of a consumable product forwards to the diagnostic test provider a fee every time the diagnostic/prognostic test provider forwards information to the subject as set forth in step (c) above.

Compositions Comprising Osteoarthritis-Directed Molecules

[0107] Featured herein is a composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and one or more molecules specifically directed and targeted to a nucleic acid comprising a *IL1RL1* nucleotide sequence or amino acid sequence. Such directed molecules include, but are not limited to, a compound that binds to a *IL1RL1* nucleotide sequence or amino acid sequence referenced herein; a RNAi or siRNA molecule having a strand complementary or substantially complementary to a *IL1RL1* nucleotide sequence (e.g., hybridizes to a *IL1RL1* nucleotide sequence under conditions of high stringency); an antisense nucleic acid complementary or substantially complementary to an RNA encoded by a *IL1RL1* nucleotide sequence (e.g., hybridizes to a *IL1RL1* nucleotide sequence under conditions of high stringency); a ribozyme that hybridizes to a *IL1RL1* nucleotide sequence (e.g., hybridizes to a *IL1RL1* nucleotide sequence under conditions of high stringency); a nucleic acid aptamer that specifically binds a polypeptide encoded by *IL1RL1* nucleotide sequence; and an antibody that specifically binds to a polypeptide encoded by *IL1RL1* nucleotide sequence or binds to a nucleic acid having such a nucleotide sequence. In specific embodiments, the osteoarthritis directed molecule interacts with a nucleic acid or polypeptide variant associated with osteoarthritis, such as variants referenced herein. In other embodiments, the osteoarthritis directed molecule interacts with a polypeptide involved in a signal pathway of a polypeptide encoded by a *IL1RL1* nucleotide sequence, or a nucleic acid comprising such a nucleotide sequence.

[0108] Compositions sometimes include an adjuvant known to stimulate an immune response, and in certain embodiments, an adjuvant that stimulates a T-cell lymphocyte response. Adjuvants are known, including but not limited to an aluminum adjuvant (e.g., aluminum hydroxide); a cytokine adjuvant or adjuvant that stimulates a cytokine response (e.g., interleukin (IL)-12 and/or gamma-interferon cytokines); a Freund-type mineral oil adjuvant emulsion (e.g., Freund's complete or incomplete adjuvant); a synthetic lipoid compound; a copolymer adjuvant (e.g., TitreMax); a saponin; Quil A; a liposome; an oil-in-water emulsion (e.g., an emulsion stabilized by Tween 80 and pluronic polyoxyethlene/polyoxypropylene block copolymer (Syntex Adjuvant Formulation); TitreMax; detoxified endotoxin (MPL) and mycobacterial cell wall components (TDW, CWS) in 2% squalene (Ribi

Adjuvant System)); a muramyl dipeptide; an immune-stimulating complex (ISCOM, e.g., an Ag-modified saponin/cholesterol micelle that forms stable cage-like structure); an aqueous phase adjuvant that does not have a depot effect (e.g., Gerbu adjuvant); a carbohydrate polymer (e.g., AdjuPrime); L-tyrosine; a manide-oleate compound (e.g., Montanide); an ethylene-vinyl acetate copolymer (e.g., Elvax 40W1,2); or lipid A, for example. Such compositions are useful for generating an immune response against osteoarthritis directed molecule (e.g., an HLA-binding subsequence within a polypeptide encoded by a *IL1RL1* nucleotide sequence). In such methods, a peptide having an amino acid subsequence of a polypeptide encoded by a *IL1RL1* nucleotide sequence is delivered to a subject, where the subsequence binds to an HLA molecule and induces a CTL lymphocyte response. The peptide sometimes is delivered to the subject as an isolated peptide or as a minigene in a plasmid that encodes the peptide. Methods for identifying HLA-binding subsequences in such polypeptides are known (see e.g., publication WO02/20616 and PCT application US98/01373 for methods of identifying such sequences).

[0109] The cell may be in a group of cells cultured *in vitro* or in a tissue maintained *in vitro* or present in an animal *in vivo* (e.g., a rat, mouse, ape or human). In certain embodiments, a composition comprises a component from a cell such as a nucleic acid molecule (e.g., genomic DNA), a protein mixture or isolated protein, for example. The aforementioned compositions have utility in diagnostic, prognostic and pharmacogenomic methods described previously and in therapeutics described hereafter. Certain osteoarthritis directed molecules are described in greater detail below.

Compounds

[0110] Compounds can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive (see, e.g., Zuckermann et al., J. Med. Chem.37: 2678-85 (1994)); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; "one-bead one-compound" library methods; and synthetic library methods using affinity chromatography selection. Biological library and peptoid library approaches are typically limited to peptide libraries, while the other approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, Anticancer Drug Des. 12: 145, (1997)). Examples of methods for synthesizing molecular libraries are described, for example, in DeWitt et al., Proc. Natl. Acad. Sci. U.S.A. 90: 6909 (1993); Erb et al., Proc. Natl. Acad. Sci. USA 91: 11422 (1994); Zuckermann et al., J. Med. Chem. 37: 2678 (1994); Cho et al., Science 261: 1303 (1993); Carrell et al., Angew. Chem. Int. Ed. Engl. 33: 2059 (1994); Carell et al., Angew. Chem. Int. Ed. Engl. 33: 2061 (1994); and in Gallop et al., J. Med. Chem. 37: 1233 (1994).

[0111] Libraries of compounds may be presented in solution (e.g., Houghten, Biotechniques 13: 412-421 (1992)), or on beads (Lam, Nature 354: 82-84 (1991)), chips (Fodor, Nature 364: 555-556 (1993)), bacteria or spores (Ladner, United States Patent No. 5,223,409), plasmids (Cull et al., Proc. Natl. Acad. Sci. USA 89: 1865-1869 (1992)) or on phage (Scott and Smith, Science 249: 386-390 (1990); Devlin, Science 249: 404-406 (1990); Cwirla et al., Proc. Natl. Acad. Sci. 87: 6378-6382 (1990); Felici, J. Mol. Biol. 222: 301-310 (1991); Ladner supra.).

[0112] A compound sometimes alters expression and sometimes alters activity of a polypeptide target and may be a small molecule. Small molecules include, but are not limited to, peptides, peptidomimetics (e.g., peptoids), amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e., including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

Antisense Nucleic Acid Molecules, Ribozymes, RNAi, siRNA and Modified Nucleic Acid Molecules

[0113] An "antisense" nucleic acid refers to a nucleotide sequence complementary to a "sense" nucleic acid encoding a polypeptide, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. The antisense nucleic acid can be complementary to an entire coding strand, or to a portion thereof or a substantially identical sequence thereof. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence (e.g., 5' and 3' untranslated regions in SEQ ID NO: 1).

[0114] An antisense nucleic acid can be designed such that it is complementary to the entire coding region of an mRNA encoded by a nucleotide sequence (e.g., SEQ ID NO: 1), and often the antisense nucleic acid is an oligonucleotide antisense to only a portion of a coding or noncoding region of the mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of the mRNA, e.g., between the -10 and +10 regions of the target gene nucleotide sequence of interest. An antisense oligonucleotide can be, for example, about 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or more nucleotides in length. The antisense nucleic acids, which include the ribozymes described hereafter, can be designed to target a *IL1RL1* nucleotide sequence, often a variant associated with osteoarthritis, or a substantially identical sequence thereof. Among the variants, minor alleles and major alleles can be targeted, and those associated with a higher risk of osteoarthritis are often designed, tested, and administered to subjects.

[0115] An antisense nucleic acid can be constructed using chemical synthesis and enzymatic ligation reactions using standard procedures. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Antisense nucleic acid also can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

[0116] When utilized as therapeutics, antisense nucleic acids typically are administered to a subject (e.g., by direct injection at a tissue site) or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a polypeptide and thereby inhibit expression of the polypeptide, for example, by inhibiting transcription and/or translation. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then are administered systemically. For systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, for example, by linking antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. Antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. Sufficient intracellular concentrations of antisense molecules are achieved by incorporating a strong promoter, such as a pol II or pol III promoter, in the vector construct.

[0117] Antisense nucleic acid molecules sometimes are alpha-anomeric nucleic acid molecules. An alpha-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual beta-units, the strands run parallel to each other (Gaultier et al., Nucleic Acids. Res. 15: 6625-6641 (1987)). Antisense nucleic acid molecules can also comprise a 2'-o-methylribonucleotide (Inoue et al., Nucleic Acids Res. 15: 6131-6148 (1987)) or a chimeric RNA-DNA analogue (Inoue et al., FEBS Lett. 215: 327-330 (1987)). Antisense nucleic acids sometimes are composed of DNA or PNA or any other nucleic acid derivatives described previously.

[0118] In another embodiment, an antisense nucleic acid is a ribozyme. A ribozyme having specificity for a *IL1RL1* nucleotide sequence can include one or more sequences complementary to such a nucleotide sequence, and a sequence having a known catalytic region responsible for mRNA cleavage (see e.g., U.S. Pat. No. 5,093,246 or Haselhoff and Gerlach, Nature 334: 585-591 (1988)). For example, a derivative of a Tetrahymena L-19 IVS RNA is sometimes utilized in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a mRNA (see e.g., Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742). Also, target mRNA sequences

can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (see e.g., Bartel & Szostak, Science 261: 1411-1418 (1993)).

[0119] Osteoarthritis directed molecules include in certain embodiments nucleic acids that can form triple helix structures with a *IL1RL1* nucleotide sequence, or a substantially identical sequence thereof, especially one that includes a regulatory region that controls expression of a polypeptide. Gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of a nucleotide sequence referenced herein or a substantially identical sequence (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of a gene in target cells (see e.g., Helene, Anticancer Drug Des. 6(6): 569-84 (1991); Helene et al., Ann. N.Y. Acad. Sci. 660: 27-36 (1992); and Maher, Bioassays 14(12): 807-15 (1992). Potential sequences that can be targeted for triple helix formation can be increased by creating a so-called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

[0120] Osteoarthritis directed molecules include RNAi and siRNA nucleic acids. Gene expression may be inhibited by the introduction of double-stranded RNA (dsRNA), which induces potent and specific gene silencing, a phenomenon called RNA interference or RNAi. See, e.g., Fire et al., US Patent Number 6,506,559; Tuschl et al. PCT International Publication No. WO 01/75164; Kay et al. PCT International Publication No. WO 03/010180A1; or Bosher JM, Labouesse, Nat Cell Biol 2000 Feb;2(2):E31-6. This process has been improved by decreasing the size of the double-stranded RNA to 20-24 base pairs (to create small-interfering RNAs or siRNAs) that "switched off" genes in mammalian cells without initiating an acute phase response, i.e., a host defense mechanism that often results in cell death (see, e.g., Caplen et al. Proc Natl Acad Sci U S A. 2001 Aug 14;98(17):9742-7 and Elbashir et al. Methods 2002 Feb;26(2):199-213). There is increasing evidence of post-transcriptional gene silencing by RNA interference (RNAi) for inhibiting targeted expression in mammalian cells at the mRNA level, in human cells. There is additional evidence of effective methods for inhibiting the proliferation and migration of tumor cells in human patients, and for inhibiting metastatic cancer development (see, e.g., U.S. Patent Application No. US2001000993183; Caplen et al. Proc Natl Acad Sci U S A; and Abderrahmani et al. Mol Cell Biol 2001 Nov21(21):7256-67).

[0121] An "siRNA" or "RNAi" refers to a nucleic acid that forms a double stranded RNA and has the ability to reduce or inhibit expression of a gene or target gene when the siRNA is delivered to or expressed in the same cell as the gene or target gene. "siRNA" refers to short double-stranded RNA formed by the complementary strands. Complementary portions of the siRNA that hybridize to form the double stranded molecule often have substantial or complete identity to the target molecule sequence. In

34

one embodiment, an siRNA refers to a nucleic acid that has substantial or complete identity to a target gene and forms a double stranded siRNA.

[0122] When designing the siRNA molecules, the targeted region often is selected from a given DNA sequence beginning 50 to 100 nucleotides downstream of the start codon. See, e.g., Elbashir et al,. Methods 26:199-213 (2002). Initially, 5' or 3' UTRs and regions nearby the start codon were avoided assuming that UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP or RISC endonuclease complex. Sometimes regions of the target 23 nucleotides in length conforming to the sequence motif AA(N19)TT (N, an nucleotide), and regions with approximately 30% to 70% G/C-content (often about 50% G/C-content) often are selected. If no suitable sequences are found, the search often is extended using the motif NA(N21). The sequence of the sense siRNA sometimes corresponds to (N19) TT or N21 (position 3 to 23 of the 23-nt motif), respectively. In the latter case, the 3' end of the sense siRNA often is converted to TT. The rationale for this sequence conversion is to generate a symmetric duplex with respect to the sequence composition of the sense and antisense 3' overhangs. The antisense siRNA is synthesized as the complement to position 1 to 21 of the 23-nt motif. Because position 1 of the 23-nt motif is not recognized sequence-specifically by the antisense siRNA, the 3'-most nucleotide residue of the antisense siRNA can be chosen deliberately. However, the penultimate nucleotide of the antisense siRNA (complementary to position 2 of the 23-nt motif) often is complementary to the targeted sequence. For simplifying chemical synthesis, TT often is utilized. siRNAs corresponding to the target motif NAR(N17)YNN, where R is purine (A,G) and Y is pyrimidine (C,U), often are selected. Respective 21 nucleotide sense and antisense siRNAs often begin with a purine nucleotide and can also be expressed from pol III expression vectors without a change in targeting site. Expression of RNAs from pol III promoters often is efficient when the first transcribed nucleotide is a purine.

[0123] The sequence of the siRNA can correspond to the full length target gene, or a subsequence thereof. Often, the siRNA is about 15 to about 50 nucleotides in length (e.g., each complementary sequence of the double stranded siRNA is 15-50 nucleotides in length, and the double stranded siRNA is about 15-50 base pairs in length, sometimes about 20-30 nucleotides in length or about 20-25 nucleotides in length, e.g., 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in length. The siRNA sometimes is about 21 nucleotides in length. Methods of using siRNA are well known in the art, and specific siRNA molecules may be purchased from a number of companies including Dharmacon Research, Inc.

[0124] Antisense, ribozyme, RNAi and siRNA nucleic acids can be altered to form modified nucleic acid molecules. The nucleic acids can be altered at base moieties, sugar moieties or phosphate backbone moieties to improve stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup et al., Bioorganic & Medicinal Chemistry 4 (1): 5-23 (1996)). As used herein, the terms "peptide

35

nucleic acid" or "PNA" refers to a nucleic acid mimic such as a DNA mimic, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of a PNA can allow for specific hybridization to DNA and RNA under conditions of low ionic strength. Synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described, for example, in Hyrup et al., (1996) supra and Perry-O'Keefe et al., Proc. Natl. Acad. Sci. 93: 14670-675 (1996).

[0125] PNA nucleic acids can be used in prognostic, diagnostic, and therapeutic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNA nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (e.g., by PNA-directed PCR clamping); as "artificial restriction enzymes" when used in combination with other enzymes, (e.g., S1 nucleases (Hyrup (1996) supra)); or as probes or primers for DNA sequencing or hybridization (Hyrup et al., (1996) supra; Perry-O'Keefe supra).

[0126] In other embodiments, oligonucleotides may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across cell membranes (see e.g., Letsinger et al., Proc. Natl. Acad. Sci. USA 86: 6553-6556 (1989); Lemaitre et al., Proc. Natl. Acad. Sci. USA 84: 648-652 (1987); PCT Publication No. WO88/09810) or the blood-brain barrier (see, e.g., PCT Publication No. WO89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, e.g., Krol et al., Bio-Techniques 6: 958-976 (1988)) or intercalating agents. (See, e.g., Zon, Pharm. Res. 5: 539-549 (1988)). To this end, the oligonucleotide may be conjugated to another molecule, (e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

[0127] Also included herein are molecular beacon oligonucleotide primer and probe molecules having one or more regions complementary to a *IL1RL1* nucleotide sequence, or a substantially identical sequence thereof, two complementary regions one having a fluorophore and one a quencher such that the molecular beacon is useful for quantifying the presence of the nucleic acid in a sample. Molecular beacon nucleic acids are described, for example, in Lizardi et al., U.S. Patent No. 5,854,033; Nazarenko et al., U.S. Patent No. 5,866,336, and Livak et al., U.S. Patent 5,876,930.

Antibodies

[0128] The term "antibody" as used herein refers to an immunoglobulin molecule or immunologically active portion thereof, i.e., an antigen-binding portion. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')$_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. An antibody sometimes is a polyclonal, monoclonal, recombinant (e.g., a chimeric or humanized), fully human, non-human (e.g.,

murine), or a single chain antibody. An antibody may have effector function and can fix complement, and is sometimes coupled to a toxin or imaging agent.

[0129] A full-length polypeptide or antigenic peptide fragment encoded by a nucleotide sequence referenced herein can be used as an immunogen or can be used to identify antibodies made with other immunogens, e.g., cells, membrane preparations, and the like. An antigenic peptide often includes at least 8 amino acid residues of the amino acid sequences encoded by a nucleotide sequence referenced herein, or substantially identical sequence thereof, and encompasses an epitope. Antigenic peptides sometimes include 10 or more amino acids, 15 or more amino acids, 20 or more amino acids, or 30 or more amino acids. Hydrophilic and hydrophobic fragments of polypeptides sometimes are used as immunogens.

[0130] Epitopes encompassed by the antigenic peptide are regions located on the surface of the polypeptide (e.g., hydrophilic regions) as well as regions with high antigenicity. For example, an Emini surface probability analysis of the human polypeptide sequence can be used to indicate the regions that have a particularly high probability of being localized to the surface of the polypeptide and are thus likely to constitute surface residues useful for targeting antibody production. The antibody may bind an epitope on any domain or region on polypeptides described herein.

[0131] Also, chimeric, humanized, and completely human antibodies are useful for applications which include repeated administration to subjects. Chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, can be made using standard recombinant DNA techniques. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson et al International Application No. PCT/US86/02269; Akira, et al European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison et al European Patent Application 173,494; Neuberger et al PCT International Publication No. WO 86/01533; Cabilly et al U.S. Patent No. 4,816,567; Cabilly et al European Patent Application 125,023; Better et al., Science 240: 1041-1043 (1988); Liu et al., Proc. Natl. Acad. Sci. USA 84: 3439-3443 (1987); Liu et al., J. Immunol. 139: 3521-3526 (1987); Sun et al., Proc. Natl. Acad. Sci. USA 84: 214-218 (1987); Nishimura et al., Canc. Res. 47: 999-1005 (1987); Wood et al., Nature 314: 446-449 (1985); and Shaw et al., J. Natl. Cancer Inst. 80: 1553-1559 (1988); Morrison, S. L., Science 229: 1202-1207 (1985); Oi et al., BioTechniques 4: 214 (1986); Winter U.S. Patent 5,225,539; Jones et al., Nature 321: 552-525 (1986); Verhoeyan et al., Science 239: 1534; and Beidler et al., J. Immunol. 141: 4053-4060 (1988).

[0132] Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced using transgenic mice that are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. See, for example, Lonberg and Huszar, Int. Rev. Immunol. 13: 65-93 (1995); and U.S.

Patent Nos. 5,625,126; 5,633,425; 5,569,825; 5,661,016; and 5,545,806. In addition, companies such as Abgenix, Inc. (Fremont, CA) and Medarex, Inc. (Princeton, NJ), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above. Completely human antibodies that recognize a selected epitope also can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody (e.g., a murine antibody) is used to guide the selection of a completely human antibody recognizing the same epitope. This technology is described for example by Jespers et al., Bio/Technology 12: 899-903 (1994).

[0133] An antibody can be a single chain antibody. A single chain antibody (scFV) can be engineered (see, e.g., Colcher et al., Ann. N Y Acad. Sci. 880: 263-80 (1999); and Reiter, Clin. Cancer Res. 2: 245-52 (1996)). Single chain antibodies can be dimerized or multimerized to generate multivalent antibodies having specificities for different epitopes of the same target polypeptide.

[0134] Antibodies also may be selected or modified so that they exhibit reduced or no ability to bind an Fc receptor. For example, an antibody may be an isotype or subtype, fragment or other mutant, which does not support binding to an Fc receptor (e.g., it has a mutagenized or deleted Fc receptor binding region).

[0135] Also, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1 dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thiotepa chlorambucil, melphalan, carmustine (BCNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

[0136] Antibody conjugates can be used for modifying a given biological response. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a polypeptide such as tumor necrosis factor, gamma-interferon, alpha-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or

other growth factors. Also, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980, for example.

[0137] An antibody (e.g., monoclonal antibody) can be used to isolate target polypeptides by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, an antibody can be used to detect a target polypeptide (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the polypeptide. Antibodies can be used diagnostically to monitor polypeptide levels in tissue as part of a clinical testing procedure, e.g., to determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance (i.e., antibody labeling). Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H. Also, an antibody can be utilized as a test molecule for determining whether it can treat osteoarthritis, and as a therapeutic for administration to a subject for treating osteoarthritis.

[0138] An antibody can be made by immunizing with a purified antigen, or a fragment thereof, e.g., a fragment described herein, a membrane associated antigen, tissues, e.g., crude tissue preparations, whole cells, preferably living cells, lysed cells, or cell fractions.

[0139] Included herein are antibodies which bind only a native polypeptide, only denatured or otherwise non-native polypeptide, or which bind both, as well as those having linear or conformational epitopes. Conformational epitopes sometimes can be identified by selecting antibodies that bind to native but not denatured polypeptide. Also featured are antibodies that specifically bind to a polypeptide variant associated with osteoarthritis.

Methods for Identifying Candidate Therapeutics for Treating Osteoarthritis

[0140] Current therapies for the treatment of osteoarthritis have limited efficacy, limited tolerability and significant mechanism-based side effects, and few of the available therapies adequately address underlying defects. Current therapeutic approaches were largely developed in the absence of defined molecular targets or even a solid understanding of disease pathogenesis. Therefore, provided are methods of identifying candidate therapeutics that target biochemical pathways related to the development of osteoarthritis.

[0141] Thus, featured herein are methods for identifying a candidate therapeutic for treating osteoarthritis. The methods comprise contacting a test molecule with a target molecule in a system. A "target molecule" as used herein refers to a *IL1RL1* nucleic acid, a substantially identical nucleic acid thereof, or a fragment thereof, and an encoded polypeptide of the foregoing. The methods also comprise determining the presence or absence of an interaction between the test molecule and the target molecule, where the presence of an interaction between the test molecule and the nucleic acid or polypeptide identifies the test molecule as a candidate osteoarthritis therapeutic. The interaction between the test molecule and the target molecule may be quantified.

[0142] Test molecules and candidate therapeutics include, but are not limited to, compounds, antisense nucleic acids, siRNA molecules, ribozymes, polypeptides or proteins encoded by a *IL1RL1* nucleotide sequence, or a substantially identical sequence or fragment thereof, and immunotherapeutics (e.g., antibodies and HLA-presented polypeptide fragments). A test molecule or candidate therapeutic may act as a modulator of target molecule concentration or target molecule function in a system. A "modulator" may agonize (i.e., up-regulates) or antagonize (i.e., down-regulates) a target molecule concentration partially or completely in a system by affecting such cellular functions as DNA replication and/or DNA processing (e.g., DNA methylation or DNA repair), RNA transcription and/or RNA processing (e.g., removal of intronic sequences and/or translocation of spliced mRNA from the nucleus), polypeptide production (e.g., translation of the polypeptide from mRNA), and/or polypeptide post-translational modification (e.g., glycosylation, phosphorylation, and proteolysis of pro-polypeptides). A modulator may also agonize or antagonize a biological function of a target molecule partially or completely, where the function may include adopting a certain structural conformation, interacting with one or more binding partners, ligand binding, catalysis (e.g., phosphorylation, dephosphorylation, hydrolysis, methylation, and isomerization), and an effect upon a cellular event (*e.g.*, effecting progression of osteoarthritis).

[0143] As used herein, the term "system" refers to a cell free *in vitro* environment and a cell-based environment such as a collection of cells, a tissue, an organ, or an organism. A system is "contacted" with a test molecule in a variety of manners, including adding molecules in solution and allowing them to interact with one another by diffusion, cell injection, and any administration routes in an animal. As used herein, the term "interaction" refers to an effect of a test molecule on test molecule, where the effect sometimes is binding between the test molecule and the target molecule, and sometimes is an observable change in cells, tissue, or organism.

[0144] There are many standard methods for detecting the presence or absence of interaction between a test molecule and a target molecule. For example, titrametric, acidimetric, radiometric, NMR, monolayer, polarographic, spectrophotometric, fluorescent, and ESR assays probative of a target molecule interaction may be utilized.

40

[0145] Test molecule/target molecule interactions can be detected and/or quantified using assays known in the art. For example, an interaction can be determined by labeling the test molecule and/or the target molecule, where the label is covalently or non-covalently attached to the test molecule or target molecule. The label is sometimes a radioactive molecule such as ^{125}I, ^{131}I, ^{35}S or ^{3}H, which can be detected by direct counting of radioemission or by scintillation counting. Also, enzymatic labels such as horseradish peroxidase, alkaline phosphatase, or luciferase may be utilized where the enzymatic label can be detected by determining conversion of an appropriate substrate to product. In addition, presence or absence of an interaction can be determined without labeling. For example, a microphysiometer (e.g., Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indication of an interaction between a test molecule and target molecule (McConnell, H. M. et al., Science 257: 1906-1912 (1992)).

[0146] In cell-based systems, cells typically include a *IL1RL1* nucleic acid, an encoded polypeptide, or substantially identical nucleic acid or polypeptide thereof, and are often of mammalian origin, although the cell can be of any origin. Whole cells, cell homogenates, and cell fractions (e.g., cell membrane fractions) can be subjected to analysis. Where interactions between a test molecule with a target polypeptide are monitored, soluble and/or membrane bound forms of the polypeptide may be utilized. Where membrane-bound forms of the polypeptide are used, it may be desirable to utilize a solubilizing agent. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)$_n$, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate.

[0147] An interaction between a test molecule and target molecule also can be detected by monitoring fluorescence energy transfer (FET) (see, e.g., Lakowicz et al., U.S. Patent No. 5,631,169; Stavrianopoulos et al. U.S. Patent No. 4,868,103). A fluorophore label on a first, "donor" molecule is selected such that its emitted fluorescent energy will be absorbed by a fluorescent label on a second, "acceptor" molecule, which in turn is able to fluoresce due to the absorbed energy. Alternately, the "donor" polypeptide molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the "acceptor" molecule label may be differentiated from that of the "donor". Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, the spatial relationship between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the

41

"acceptor" molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (*e.g.*, using a fluorimeter).

[0148] In another embodiment, determining the presence or absence of an interaction between a test molecule and a target molecule can be effected by monitoring surface plasmon resonance (*see, e.g.*, Sjolander & Urbaniczk, *Anal. Chem. 63:* 2338-2345 (1991) and Szabo *et al., Curr. Opin. Struct. Biol. 5:* 699-705 (1995)). "Surface plasmon resonance" or "biomolecular interaction analysis (BIA)" can be utilized to detect biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the mass at the binding surface (indicative of a binding event) result in alterations of the refractive index of light near the surface (the optical phenomenon of surface plasmon resonance (SPR)), resulting in a detectable signal which can be used as an indication of real-time reactions between biological molecules.

[0149] In another embodiment, the target molecule or test molecules are anchored to a solid phase, facilitating the detection of target molecule/test molecule complexes and separation of the complexes from free, uncomplexed molecules. The target molecule or test molecule is immobilized to the solid support. In an embodiment, the target molecule is anchored to a solid surface, and the test molecule, which is not anchored, can be labeled, either directly or indirectly, with detectable labels discussed herein.

[0150] It may be desirable to immobilize a target molecule, an anti-target molecule antibody, and/or test molecules to facilitate separation of target molecule/test molecule complexes from uncomplexed forms, as well as to accommodate automation of the assay. The attachment between a test molecule and/or target molecule and the solid support may be covalent or non-covalent (*see, e.g.*, U.S. Patent No. 6,022,688 for non-covalent attachments). The solid support may be one or more surfaces of the system, such as one or more surfaces in each well of a microtiter plate, a surface of a silicon wafer, a surface of a bead (*see, e.g.*, Lam, *Nature 354:* 82-84 (1991)) that is optionally linked to another solid support, or a channel in a microfluidic device, for example. Types of solid supports, linker molecules for covalent and non-covalent attachments to solid supports, and methods for immobilizing nucleic acids and other molecules to solid supports are well known (*see, e.g.*, U.S. Patent Nos. 6,261,776; 5,900,481; 6,133,436; and 6,022,688; and WIPO publication WO 01/18234).

[0151] In an embodiment, target molecule may be immobilized to surfaces via biotin and streptavidin. For example, biotinylated target polypeptide can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). In another embodiment, a target polypeptide can be prepared as a fusion polypeptide. For example, glutathione-S-transferase/target polypeptide fusion can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivitized microtiter plates, which are then combined with a test molecule

under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, or the matrix is immobilized in the case of beads, and complex formation is determined directly or indirectly as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of target molecule binding or activity is determined using standard techniques.

[0152] In an embodiment, the non-immobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) under conditions such that a significant percentage of complexes formed will remain immobilized to the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of manners. Where the previously non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously non-immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface, *e.g.*, by adding a labeled antibody specific for the immobilized component, where the antibody, in turn, can be directly labeled or indirectly labeled with, *e.g.*, a labeled anti-Ig antibody.

[0153] In another embodiment, an assay is performed utilizing antibodies that specifically bind target molecule or test molecule but do not interfere with binding of the target molecule to the test molecule. Such antibodies can be derivitized to a solid support, and unbound target molecule may be immobilized by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule.

[0154] Cell free assays also can be conducted in a liquid phase. In such an assay, reaction products are separated from unreacted components, by any of a number of standard techniques, including but not limited to: differential centrifugation (*see, e.g.*, Rivas, G., and Minton, *Trends Biochem Sci Aug;18(8):* 284-7 (1993)); chromatography (gel filtration chromatography, ion-exchange chromatography); electrophoresis (*see, e.g.*, Ausubel *et al., eds. Current Protocols in Molecular Biology*, J. Wiley: New York (1999)); and immunoprecipitation (*see, e.g.*, Ausubel *et al., eds., supra*). Media and chromatographic techniques are known to one skilled in the art (*see, e.g.*, Heegaard, *J Mol. Recognit. Winter; 11(1-6):* 141-8 (1998); Hage & Tweed, *J. Chromatogr. B Biomed. Sci. Appl. Oct 10; 699 (1-2): 499-525* (1997)). Further, fluorescence energy transfer may also be conveniently utilized, as described herein, to detect binding without further purification of the complex from solution.

[0155] In another embodiment, modulators of target molecule expression are identified. For example, a cell or cell free mixture is contacted with a candidate compound and the expression of target mRNA or target polypeptide is evaluated relative to the level of expression of target mRNA or target

polypeptide in the absence of the candidate compound. When expression of target mRNA or target polypeptide is greater in the presence of the candidate compound than in its absence, the candidate compound is identified as an agonist of target mRNA or target polypeptide expression. Alternatively, when expression of target mRNA or target polypeptide is less (*e.g.*, less with statistical significance) in the presence of the candidate compound than in its absence, the candidate compound is identified as an antagonist or inhibitor of target mRNA or target polypeptide expression. The level of target mRNA or target polypeptide expression can be determined by methods described herein.

[0156] In another embodiment, binding partners that interact with a target molecule are detected. The target molecules can interact with one or more cellular or extracellular macromolecules, such as polypeptides *in vivo*, and these interacting molecules are referred to herein as "binding partners." Binding partners can agonize or antagonize target molecule biological activity. Also, test molecules that agonize or antagonize interactions between target molecules and binding partners can be useful as therapeutic molecules as they can up-regulate or down-regulated target molecule activity *in vivo* and thereby treat osteoarthritis.

[0157] Binding partners of target molecules can be identified by methods known in the art. For example, binding partners may be identified by lysing cells and analyzing cell lysates by electrophoretic techniques. Alternatively, a two-hybrid assay or three-hybrid assay can be utilized (*see, e.g.*, U.S. Patent No. 5,283,317; Zervos *et al., Cell 72:223-232* (1993); Madura *et al., J. Biol. Chem. 268:* 12046-12054 (1993); Bartel *et al., Biotechniques 14:* 920-924 (1993); Iwabuchi *et al., Oncogene 8:* 1693-1696 (1993); and Brent WO94/10300). A two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. The assay often utilizes two different DNA constructs. In one construct, a *IL1RL1* nucleic acid (sometimes referred to as the "bait") is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In another construct, a DNA sequence from a library of DNA sequences that encodes a potential binding partner (sometimes referred to as the "prey") is fused to a gene that encodes an activation domain of the known transcription factor. Sometimes, a *IL1RL1* nucleic acid can be fused to the activation domain. If the "bait" and the "prey" molecules interact *in vivo,* the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to identify the potential binding partner.

[0158] In an embodiment for identifying test molecules that antagonize or agonize complex formation between target molecules and binding partners, a reaction mixture containing the target molecule and the binding partner is prepared, under conditions and for a time sufficient to allow complex formation. The reaction mixture often is provided in the presence or absence of the test molecule. The

test molecule can be included initially in the reaction mixture, or can be added at a time subsequent to the addition of the target molecule and its binding partner. Control reaction mixtures are incubated without the test molecule or with a placebo. Formation of any complexes between the target molecule and the binding partner then is detected. Decreased formation of a complex in the reaction mixture containing test molecule as compared to in a control reaction mixture indicates that the molecule antagonizes target molecule/binding partner complex formation. Alternatively, increased formation of a complex in the reaction mixture containing test molecule as compared to in a control reaction mixture indicates that the molecule agonizes target molecule/binding partner complex formation. In another embodiment, complex formation of target molecule/binding partner can be compared to complex formation of mutant target molecule/binding partner (*e.g.*, amino acid modifications in a target polypeptide). Such a comparison can be important in those cases where it is desirable to identify test molecules that modulate interactions of mutant but not non-mutated target gene products.

[0159] The assays can be conducted in a heterogeneous or homogeneous format. In heterogeneous assays, target molecule and/or the binding partner are immobilized to a solid phase, and complexes are detected on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the molecules being tested. For example, test compounds that agonize target molecule/binding partner interactions can be identified by conducting the reaction in the presence of the test molecule in a competition format. Alternatively, test molecules that agonize preformed complexes, *e.g.*, molecules with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed.

[0160] In a heterogeneous assay embodiment, the target molecule or the binding partner is anchored onto a solid surface (*e.g.*, a microtiter plate), while the non-anchored species is labeled, either directly or indirectly. The anchored molecule can be immobilized by non-covalent or covalent attachments. Alternatively, an immobilized antibody specific for the molecule to be anchored can be used to anchor the molecule to the solid surface. The partner of the immobilized species is exposed to the coated surface with or without the test molecule. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) such that a significant portion of any complexes formed will remain immobilized on the solid surface. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface is indicative of complex. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored to the surface; *e.g.*, by using a labeled antibody specific for the initially non-immobilized species. Depending upon the order of addition of reaction components, test compounds that inhibit complex formation or that disrupt preformed complexes can be detected.

[0161] In another embodiment, the reaction can be conducted in a liquid phase in the presence or absence of test molecule, where the reaction products are separated from unreacted components, and the complexes are detected (e.g., using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes). Again, depending upon the order of addition of reactants to the liquid phase, test compounds that inhibit complex or that disrupt preformed complexes can be identified.

[0162] In an alternate embodiment, a homogeneous assay can be utilized. For example, a preformed complex of the target gene product and the interactive cellular or extracellular binding partner product is prepared. One or both of the target molecule or binding partner is labeled, and the signal generated by the label(s) is quenched upon complex formation (e.g., U.S. Patent No. 4,109,496 that utilizes this approach for immunoassays). Addition of a test molecule that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances that disrupt target molecule/binding partner complexes can be identified.

[0163] Candidate therapeutics for treating osteoarthritis are identified from a group of test molecules that interact with a target molecule. Test molecules are normally ranked according to the degree with which they modulate (e.g., agonize or antagonize) a function associated with the target molecule (e.g., DNA replication and/or processing, RNA transcription and/or processing, polypeptide production and/or processing, and/or biological function/activity), and then top ranking modulators are selected. Also, pharmacogenomic information described herein can determine the rank of a modulator. The top 10% of ranked test molecules often are selected for further testing as candidate therapeutics, and sometimes the top 15%, 20%, or 25% of ranked test molecules are selected for further testing as candidate therapeutics. Candidate therapeutics typically are formulated for administration to a subject.

Therapeutic Formulations

[0164] Formulations and pharmaceutical compositions typically include in combination with a pharmaceutically acceptable carrier one or more target molecule modulators. The modulator often is a test molecule identified as having an interaction with a target molecule by a screening method described above. The modulator may be a compound, an antisense nucleic acid, a ribozyme, an antibody, or a binding partner. Also, formulations may comprise a target polypeptide or fragment thereof in combination with a pharmaceutically acceptable carrier.

[0165] As used herein, the term "pharmaceutically acceptable carrier" includes solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into the compositions. Pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

[0166] A pharmaceutical composition typically is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerin, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

[0167] Oral compositions generally include an inert diluent or an edible carrier. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules, e.g., gelatin capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

[0168] Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It should be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sodium chloride in

47

the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

[0169] Sterile injectable solutions can be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[0170] For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

[0171] Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art. Molecules can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

[0172] In one embodiment, active molecules are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. Materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

[0173] It is advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity

48

of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier.

[0174] Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Molecules which exhibit high therapeutic indices are preferred. While molecules that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

[0175] The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such molecules lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any molecules used in the methods described herein, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

[0176] As defined herein, a therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, sometimes about 0.01 to 25 mg/kg body weight, often about 0.1 to 20 mg/kg body weight, and more often about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The protein or polypeptide can be administered one time per week for between about 1 to 10 weeks, sometimes between 2 to 8 weeks, often between about 3 to 7 weeks, and more often for about 4, 5, or 6 weeks. The skilled artisan will appreciate that certain factors may influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

[0177] With regard to polypeptide formulations, featured herein is a method for treating osteoarthritis in a subject, which comprises contacting one or more cells in the subject with a first polypeptide, where the subject comprises a second polypeptide having one or more polymorphic variations associated with cancer, and where the first polypeptide comprises fewer polymorphic

variations associated with cancer than the second polypeptide. The first and second polypeptides are encoded by a nucleic acid which comprises a nucleotide sequence in SEQ ID NO: 1-4; a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence referenced in SEQ ID NO: 1-4; a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-4 and a nucleotide sequence 90% or more identical to a nucleotide sequence in SEQ ID NO: 1-4. The subject often is a human.

[0178] For antibodies, a dosage of 0.1 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg) is often utilized. If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is often appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (*e.g.*, into the brain). A method for lipidation of antibodies is described by Cruikshank *et al.*, *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* *14*:193 (1997).

[0179] Antibody conjugates can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a polypeptide such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors. Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

[0180] For compounds, exemplary doses include milligram or microgram amounts of the compound per kilogram of subject or sample weight, for example, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. It is understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid described herein, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of

factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

[0181] With regard to nucleic acid formulations, gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (*see, e.g.,* U.S. Patent 5,328,470) or by stereotactic injection (*see e.g.,* Chen *et al.,* (1994) *Proc. Natl. Acad. Sci. USA 91*:3054-3057). Pharmaceutical preparations of gene therapy vectors can include a gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells (*e.g.,* retroviral vectors) the pharmaceutical preparation can include one or more cells which produce the gene delivery system. Examples of gene delivery vectors are described herein.

Therapeutic Methods ·

[0182] A therapeutic formulation described above can be administered to a subject in need of a therapeutic for inducing a desired biological response.. Therapeutic formulations can be administered by any of the paths described herein. With regard to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from pharmacogenomic analyses described herein.

[0183] As used herein, the term "treatment" is defined as the application or administration of a therapeutic formulation to a subject, or application or administration of a therapeutic agent to an isolated tissue or cell line from a subject with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect osteoarthritis, symptoms of osteoarthritis or a predisposition towards osteoarthritis. A therapeutic formulation includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides. Administration of a therapeutic formulation can occur prior to the manifestation of symptoms characteristic of osteoarthritis, such that osteoarthritis is prevented or delayed in its progression. The appropriate therapeutic composition can be determined based on screening assays described herein.

[0184] As discussed, successful treatment of osteoarthritis can be brought about by techniques that serve to agonize target molecule expression or function, or alternatively, antagonize target molecule expression or function. These techniques include administration of modulators that include, but are not limited to, small organic or inorganic molecules; antibodies (including, for example, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and Fab, F(ab')$_2$ and Fab expression library fragments, scFV molecules, and epitope-binding fragments thereof); and peptides, phosphopeptides, or polypeptides.

[0185] Further, antisense and ribozyme molecules that inhibit expression of the target gene can also be used to reduce the level of target gene expression, thus effectively reducing the level of target gene activity. Still further, triple helix molecules can be utilized in reducing the level of target gene activity. Antisense, ribozyme and triple helix molecules are discussed above. It is possible that the use of antisense, ribozyme, and/or triple helix molecules to reduce or inhibit mutant gene expression can also reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles, such that the concentration of normal target gene product present can be lower than is necessary for a normal phenotype. In such cases, nucleic acid molecules that encode and express target gene polypeptides exhibiting normal target gene activity can be introduced into cells via gene therapy method. Alternatively, in instances in that the target gene encodes an extracellular polypeptide, it can be preferable to co-administer normal target gene polypeptide into the cell or tissue in order to maintain the requisite level of cellular or tissue target gene activity.

[0186] Another method by which nucleic acid molecules may be utilized in treating or preventing osteoarthritis is use of aptamer molecules specific for target molecules. Aptamers are nucleic acid molecules having a tertiary structure which permits them to specifically bind to ligands (*see, e.g.,* Osborne, *et al., Curr. Opin. Chem. Biol.1(1):* 5-9 (1997); and Patel, D. J., *Curr. Opin. Chem. Biol. Jun;1(1):* 32-46 (1997)).

[0187] Yet another method of utilizing nucleic acid molecules for osteoarthritis treatment is gene therapy, which can also be referred to as allele therapy. Provided herein is a gene therapy method for treating osteoarthritis in a subject, which comprises contacting one or more cells in the subject or from the subject with a nucleic acid having a first nucleotide sequence (e.g., the first nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1-4). Genomic DNA in the subject comprises a second nucleotide sequence having one or more polymorphic variations associated with osteoarthritis (e.g., the second nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1). The first and second nucleotide sequences typically are substantially identical to one another, and the first nucleotide sequence comprises fewer polymorphic variations associated with osteoarthritis than the second nucleotide sequence. The first nucleotide sequence may comprise a gene sequence that encodes a full-length polypeptide or a fragment thereof. The subject is often a human. Allele therapy methods often are utilized in conjunction with a method of first determining whether a subject has genomic DNA that includes polymorphic variants associated with osteoarthritis.

[0188] In another allele therapy embodiment, provided herein is a method which comprises contacting one or more cells in the subject or from the subject with a polypeptide encoded by a nucleic acid having a first nucleotide sequence (e.g., the first nucleotide sequence is identical to or substantially identical to the nucleotide sequence of SEQ ID NO: 1-4). Genomic DNA in the subject comprises a

52

second nucleotide sequence having one or more polymorphic variations associated with osteoarthritis (e.g., the second nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1). The first and second nucleotide sequences typically are substantially identical to one another, and the first nucleotide sequence comprises fewer polymorphic variations associated with osteoarthritis than the second nucleotide sequence. The first nucleotide sequence may comprise a gene sequence that encodes a full-length polypeptide or a fragment thereof. The subject is often a human.

[0189] For antibody-based therapies, antibodies can be generated that are both specific for target molecules and that reduce target molecule activity. Such antibodies may be administered in instances where antagonizing a target molecule function is appropriate for the treatment of osteoarthritis.

[0190] In circumstances where stimulating antibody production in an animal or a human subject by injection with a target molecule is harmful to the subject, it is possible to generate an immune response against the target molecule by use of anti-idiotypic antibodies (see, e.g., Herlyn, Ann. Med.;31(1): 66-78 (1999); and Bhattacharya-Chatterjee & Foon, Cancer Treat. Res.; 94: 51-68 (1998)). Introducing an anti-idiotypic antibody to a mammal or human subject often stimulates production of anti-anti-idiotypic antibodies, which typically are specific to the target molecule. Vaccines directed to osteoarthritis also may be generated in this fashion.

[0191] In instances where the target molecule is intracellular and whole antibodies are used, internalizing antibodies may be preferred. Lipofectin or liposomes can be used to deliver the antibody or a fragment of the Fab region that binds to the target antigen into cells. Where fragments of the antibody are used, the smallest inhibitory fragment that binds to the target antigen is preferred. For example, peptides having an amino acid sequence corresponding to the Fv region of the antibody can be used. Alternatively, single chain neutralizing antibodies that bind to intracellular target antigens can also be administered. Such single chain antibodies can be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population (see, e.g., Marasco et al., Proc. Natl. Acad. Sci. USA 90: 7889-7893 (1993)).

[0192] Modulators can be administered to a patient at therapeutically effective doses to treat osteoarthritis. A therapeutically effective dose refers to an amount of the modulator sufficient to result in amelioration of symptoms of osteoarthritis. Toxicity and therapeutic efficacy of modulators can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Modulators that exhibit large therapeutic indices are preferred. While modulators that exhibit toxic side effects can be used, care should be taken to design a delivery system that targets such molecules to the site of affected tissue in order to minimize potential damage to uninfected cells, thereby reducing side effects.

[0193] Data obtained from cell culture assays and animal studies can be used in formulating a range of dosages for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the methods described herein, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (*i.e.*, the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma can be measured, for example, by high performance liquid chromatography.

[0194] Another example of effective dose determination for an individual is the ability to directly assay levels of "free" and "bound" compound in the serum of the test subject. Such assays may utilize antibody mimics and/or "biosensors" that have been created through molecular imprinting techniques. Molecules that modulate target molecule activity are used as a template, or "imprinting molecule", to spatially organize polymerizable monomers prior to their polymerization with catalytic reagents. The subsequent removal of the imprinted molecule leaves a polymer matrix which contains a repeated "negative image" of the compound and is able to selectively rebind the molecule under biological assay conditions. A detailed review of this technique can be seen in Ansell *et al., Current Opinion in Biotechnology 7:* 89-94 (1996) and in Shea, *Trends in Polymer Science 2:* 166-173 (1994). Such "imprinted" affinity matrixes are amenable to ligand-binding assays, whereby the immobilized monoclonal antibody component is replaced by an appropriately imprinted matrix. An example of the use of such matrixes in this way can be seen in Vlatakis, *et al., Nature 361:* 645-647 (1993). Through the use of isotope-labeling, the "free" concentration of compound which modulates target molecule expression or activity readily can be monitored and used in calculations of IC_{50}. Such "imprinted" affinity matrixes can also be designed to include fluorescent groups whose photon-emitting properties measurably change upon local and selective binding of target compound. These changes readily can be assayed in real time using appropriate fiberoptic devices, in turn allowing the dose in a test subject to be quickly optimized based on its individual IC_{50}. An example of such a "biosensor" is discussed in Kriz *et al., Analytical Chemistry 67:* 2142-2144 (1995).

[0195] The examples set forth below are intended to illustrate but not limit the invention.

Examples

[0196] In the following studies a group of subjects was selected according to specific parameters relating to osteoarthritis. Nucleic acid samples obtained from individuals in the study group were subjected to genetic analysis, which identified associations between osteoarthritis and a polymorphism in

the *IL1RL1* gene on chromosome two. The polymorphism was genotyped again in two replication cohorts consisting of individuals selected for OA. In addition, SNPs proximal to the incident polymorphism were identified and allelotyped in OA case and control pools. Methods are described for producing *IL1RL1* polypeptide and *IL1RL1* polypeptide variants *in vitro* or *in vivo*, *IL1RL1* nucleic acids or polypeptides and variants thereof are utilized for screening test molecules for those that interact with *IL1RL1* molecules. Test molecules identified as interactors with *IL1RL1* molecules and *IL1RL1* variants are further screened *in vivo* to determine whether they treat osteoarthritis.

<div align="center">

Example 1

Samples and Pooling Strategies

</div>

Sample Selection

[0197] Blood samples were collected from individuals diagnosed with knee osteoarthritis, which were referred to as case samples. Also, blood samples were collected from individuals not diagnosed with knee osteoarthritis as gender and age-matched controls. A database was created that listed all phenotypic trait information gathered from individuals for each case and control sample. Genomic DNA was extracted from each of the blood samples for genetic analyses.

DNA Extraction from Blood Samples

[0198] Six to ten milliliters of whole blood was transferred to a 50 ml tube containing 27 ml of red cell lysis solution (RCL). The tube was inverted until the contents were mixed. Each tube was incubated for 10 minutes at room temperature and inverted once during the incubation. The tubes were then centrifuged for 20 minutes at 3000 x g and the supernatant was carefully poured off. 100-200 µl of residual liquid was left in the tube and was pipetted repeatedly to resuspend the pellet in the residual supernatant. White cell lysis solution (WCL) was added to the tube and pipetted repeatedly until completely mixed. While no incubation was normally required, the solution was incubated at 37°C or room temperature if cell clumps were visible after mixing until the solution was homogeneous. 2 ml of protein precipitation was added to the cell lysate. The mixtures were vortexed vigorously at high speed for 20 sec to mix the protein precipitation solution uniformly with the cell lysate, and then centrifuged for 10 minutes at 3000 x g. The supernatant containing the DNA was then poured into a clean 15 ml tube, which contained 7 ml of 100% isopropanol. The samples were mixed by inverting the tubes gently until white threads of DNA were visible. Samples were centrifuged for 3 minutes at 2000 x g and the DNA was visible as a small white pellet. The supernatant was decanted and 5 ml of 70% ethanol was added to each tube. Each tube was inverted several times to wash the DNA pellet, and then centrifuged for 1 minute at 2000 x g. The ethanol was decanted and each tube was drained on clean absorbent paper. The

DNA was dried in the tube by inversion for 10 minutes, and then 1000 µl of 1X TE was added. The size of each sample was estimated, and less TE buffer was added during the following DNA hydration step if the sample was smaller. The DNA was allowed to rehydrate overnight at room temperature, and DNA samples were stored at 2-8°C.

[0199] DNA was quantified by placing samples on a hematology mixer for at least 1 hour. DNA was serially diluted (typically 1:80, 1:160, 1:320, and 1:640 dilutions) so that it would be within the measurable range of standards. 125 µl of diluted DNA was transferred to a clear U-bottom microtitre plate, and 125 µl of 1X TE buffer was transferred into each well using a multichannel pipette. The DNA and 1X TE were mixed by repeated pipetting at least 15 times, and then the plates were sealed. 50 µl of diluted DNA was added to wells A5-H12 of a black flat bottom microtitre plate. Standards were inverted six times to mix them, and then 50 µl of 1X TE buffer was pipetted into well A1, 1000 ng/ml of standard was pipetted into well A2, 500 ng/ml of standard was pipetted into well A3, and 250 ng/ml of standard was pipetted into well A4. PicoGreen (Molecular Probes, Eugene, Oregon) was thawed and freshly diluted 1:200 according to the number of plates that were being measured. PicoGreen was vortexed and then 50µl was pipetted into all wells of the black plate with the diluted DNA. DNA and PicoGreen were mixed by pipetting repeatedly at least 10 times with the multichannel pipette. The plate was placed into a Fluoroskan Ascent Machine (microplate fluorometer produced by Labsystems) and the samples were allowed to incubate for 3 minutes before the machine was run using filter pairs 485 nm excitation and 538 nm emission wavelengths. Samples having measured DNA concentrations of greater than 450 ng/µl were re-measured for conformation. Samples having measured DNA concentrations of 20 ng/µl or less were re-measured for confirmation.

Pooling Strategies – Discovery Cohort

[0200] Samples were derived from the Nottingham knee OA family study (UK) where index cases were identified through a knee replacement registry. Siblings were approached and assessed with knee x-rays and assigned status as affected or unaffected. In all 1,157 individuals were available. In order to create same-sex pools of appropriate sizes, 335 unrelated female individuals with OA from the Nottingham OA sample were selected for the case pool. The control pool was made up of unrelated female individuals from the St. Thomas twin study (England) with normal knee x-rays and without other indications of OA, regardless of anatomical location, as well as lacking family history of OA. The St. Thomas twin study consists of Caucasian, female participants from the St. Thomas' Hospital, London, adult-twin registry, which is a voluntary registry of >4,000 twin pairs ranging from 18 to 76 years of age. The female case samples and female control samples are described further in Table 1 below.

[0201] A select set of samples from each group were utilized to generate pools, and one pool was created for each group. Each individual sample in a pool was represented by an equal amount of genomic DNA. For example, where 25 ng of genomic DNA was utilized in each PCR reaction and there were 200 individuals in each pool, each individual would provide 125 pg of genomic DNA. Inclusion or exclusion of samples for a pool was based upon the following criteria: the sample was derived from an individual characterized as Caucasian; the sample was derived from an individual of British paternal and maternal descent; case samples were derived from individuals diagnosed with specific knee osteoarthritis (OA) and were recruited from an OA knee replacement clinic. Control samples were derived from individuals free of OA, family history of OA, and rheumatoid arthritis. Also, sufficient genomic DNA was extracted from each blood sample for all allelotyping and genotyping reactions performed during the study. Phenotype information from each individual was collected and included age of the individual, gender, family history of OA, general medical information (e.g., height, weight, thyroid disease, diabetes, psoriasis, hysterectomy), joint history (previous and current symptoms, joint-related operations, age at onset of symptoms, date of primary diagnosis, age of individual as of primary diagnosis and order of involvement), and knee-related findings (crepitus, restricted passive movement, bony swelling/deformity). Additional knee information included knee history, current symptoms, any major knee injury, menisectomy, knee replacement surgery, age of surgery, and treatment history (including hormone replace therapy (HRT)). Samples that met these criteria were added to appropriate pools based on disease status.

[0202] The selection process yielded the pools set forth in Table 1, which were used in the studies that follow:

TABLE 1

	Female case	Female control
Pool size (Number)	335	335
Pool Criteria (ex: case/control)	control	case
Mean Age (ex: years)	57.21	69.95

Example 2

Association of Polymorphic Variants with Osteoarthritis

[0203] A whole-genome screen was performed to identify particular SNPs associated with occurrence of osteoarthritis. As described in Example 1, two sets of samples were utilized, which included samples from female individuals having knee osteoarthritis (osteoarthritis cases), and samples

57

from female individuals not having knee osteoarthritis (female controls). The initial screen of each pool was performed in an allelotyping study, in which certain samples in each group were pooled. By pooling DNA from each group, an allele frequency for each SNP in each group was calculated. These allele frequencies were then compared to one another. Particular SNPs were considered as being associated with osteoarthritis when allele frequency differences calculated between case and control pools were statistically significant. SNP disease association results obtained from the allelotyping study were then validated by genotyping each associated SNP across all samples from each pool. The results of the genotyping then were analyzed, allele frequencies for each group were calculated from the individual genotyping results, and a p-value was calculated to determine whether the case and control groups had statistically significant differences in allele frequencies for a particular SNP. When the genotyping results agreed with the original allelotyping results, the SNP disease association was considered validated at the genetic level.

<u>SNP Panel Used for Genetic Analyses</u>

[0204] A whole-genome SNP screen began with an initial screen of approximately 25,000 SNPs over each set of disease and control samples using a pooling approach. The pools studied in the screen are described in Example 1. The SNPs analyzed in this study were part of a set of 25,488 SNPs confirmed as being statistically polymorphic as each is characterized as having a minor allele frequency of greater than 10%. The SNPs in the set reside in genes or in close proximity to genes, and many reside in gene exons. Specifically, SNPs in the set are located in exons, introns, and within 5,000 base-pairs upstream of a transcription start site of a gene. In addition, SNPs were selected according to the following criteria: they are located in ESTs; they are located in Locuslink or Ensembl genes; and they are located in Genomatix promoter predictions. SNPs in the set were also selected on the basis of even spacing across the genome, as depicted in Table 2.

[0205] A case-control study design using a whole genome association strategy involving approximately 28,000 single nucleotide polymorphisms (SNPs) was employed. Approximately 25,000 SNPs were evenly spaced in gene-based regions of the human genome with a median inter-marker distance of about 40,000 base pairs. Additionally, approximately 3,000 SNPs causing amino acid substitutions in genes described in the literature as candidates for various diseases were used. The case-control study samples were of female Caucasian origin (British paternal and maternal descent) 670 individuals were equally distributed in two groups: female controls and female cases. The whole genome association approach was first conducted on 2 DNA pools representing the 2 groups. Significant markers were confirmed by individual genotyping.

TABLE 2

General Statistics		Spacing Statistics	
Total # of SNPs	25,488	Median	37,058 bp
# of Exonic SNPs	>4,335 (17%)	Minimum*	1,000 bp
# SNPs with refSNP ID	20,776 (81%)	Maximum*	3,000,000 bp
Gene Coverage	>10,000	Mean	122,412 bp
Chromosome Coverage	All	Std Deviation	373,325 bp
		*Excludes outliers	

Allelotyping and Genotyping Results

[0206] The genetic studies summarized above and described in more detail below identified an allelic variant in the IL1RL1 gene that is associated with osteoarthritis.

Assay for Verifying, Allelotyping, and Genotyping SNPs

[0207] A MassARRAY™ system (Sequenom, Inc.) was utilized to perform SNP genotyping in a high-throughput fashion. This genotyping platform was complemented by a homogeneous, single-tube assay method (hME™ or homogeneous MassEXTEND™ (Sequenom, Inc.)) in which two genotyping primers anneal to and amplify a genomic target surrounding a polymorphic site of interest. A third primer (the MassEXTEND™ primer), which is complementary to the amplified target up to but not including the polymorphism, was then enzymatically extended one or a few bases through the polymorphic site and then terminated.

[0208] For each polymorphism, SpectroDESIGNER™ software (Sequenom, Inc.) was used to generate a set of PCR primers and a MassEXTEND™ primer which where used to genotype the polymorphism. Other primer design software could be used or one of ordinary skill in the art could manually design primers based on his or her knowledge of the relevant factors and considerations in designing such primers. Table 3 shows PCR primers and Table 4 shows extension primers used for analyzing polymorphisms. The initial PCR amplification reaction was performed in a 5 μl total volume containing 1X PCR buffer with 1.5 mM MgCl₂ (Qiagen), 200 μM each of dATP, dGTP, dCTP, dTTP (Gibco-BRL), 2.5 ng of genomic DNA, 0.1 units of HotStar DNA polymerase (Qiagen), and 200 nM each of forward and reverse PCR primers specific for the polymorphic region of interest.

TABLE 3: PCR Primers

SNP Reference	Forward PCR primer	Reverse PCR primer
rs1041973	ACGTTGGATGGGGACTTCTGACAATACAGG	ACGTTGGATGAATCGTGTGTTTGCCTCAGG

[0209] Samples were incubated at 95°C for 15 minutes, followed by 45 cycles of 95°C for 20 seconds, 56°C for 30 seconds, and 72°C for 1 minute, finishing with a 3 minute final extension at 72°C. Following amplification, shrimp alkaline phosphatase (SAP) (0.3 units in a 2 μl volume) (Amersham Pharmacia) was added to each reaction (total reaction volume was 7 μl) to remove any residual dNTPs that were not consumed in the PCR step. Samples were incubated for 20 minutes at 37°C, followed by 5 minutes at 85°C to denature the SAP.

[0210] Once the SAP reaction was complete, a primer extension reaction was initiated by adding a polymorphism-specific MassEXTEND™ primer cocktail to each sample. Each MassEXTEND™ cocktail included a specific combination of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) used to distinguish polymorphic alleles from one another. Methods for verifying, allelotyping and genotyping SNPs are disclosed, for example, in U.S. Pat. No. 6,258,538, the content of which is hereby incorporated by reference. In Table 4, ddNTPs are shown and the fourth nucleotide not shown is the dNTP.

TABLE 4: Extension Primers

SNP Reference	Extend Probe	Termination Mix
rs1041973	ATACCAGAATCAGCAACT	ACT

[0211] The MassEXTEND™ reaction was performed in a total volume of 9 μl, with the addition of 1X ThermoSequenase buffer, 0.576 units of ThermoSequenase (Amersham Pharmacia), 600 nM MassEXTEND™ primer, 2 mM of ddATP and/or ddCTP and/or ddGTP and/or ddTTP, and 2 mM of dATP or dCTP or dGTP or dTTP. The deoxy nucleotide (dNTP) used in the assay normally was complementary to the nucleotide at the polymorphic site in the amplicon. Samples were incubated at 94°C for 2 minutes, followed by 55 cycles of 5 seconds at 94°C, 5 seconds at 52°C, and 5 seconds at 72°C.

[0212] Following incubation, samples were desalted by adding 16 μl of water (total reaction volume was 25 μl), 3 mg of SpectroCLEAN™ sample cleaning beads (Sequenom, Inc.) and allowed to incubate for 3 minutes with rotation. Samples were then robotically dispensed using a piezoelectric dispensing device (SpectroJET™ (Sequenom, Inc.)) onto either 96-spot or 384-spot silicon chips containing a matrix that crystallized each sample (SpectroCHIP™ (Sequenom, Inc.)). Subsequently, MALDI-TOF mass spectrometry (Biflex and Autoflex MALDI-TOF mass spectrometers (Bruker Daltonics) can be used) and SpectroTYPER RT™ software (Sequenom, Inc.) were used to analyze and interpret the SNP genotype for each sample.

Genetic Analysis

[0213] Minor allelic frequencies for the polymorphisms set forth in Table A were verified as being 10% or greater using the extension assay described above in a group of samples isolated from 92 individuals originating from the state of Utah in the United States, Venezuela and France (Coriell cell repositories).

[0214] Genotyping results are shown for female pools in Table 5. In Table 5, "AF" refers to allelic frequency; and "F case" and "F control" refer to female case and female control groups, respectively.

TABLE 5: Genotyping Results

SNP Reference	AF F case	AF F control	p-value
rs1041973	A = 0.189 C = 0.811	A = 0.233 C = 0.767	0.0539

[0215] All of the single marker alleles set forth in Table A were considered validated, since the genotyping data agreed with the allelotyping data and each SNP significantly associated with osteoarthritis. Particularly significant associations with osteoarthritis are indicated by a calculated p-value of less than 0.05 for genotype results. SNP rs1041973 has a calculated p-value of greater than 0.05 (0.0539), but was included because it is an exonic SNP located in the IL1RL1 gene, which plays a role in inflammation and is a compelling target for osteoarthritis.

Example 3

Association of Polymorphic Variants with Osteoarthritis in Replication Cohorts

[0216] The single marker polymorphism set forth in Table A was genotyped again in two replication cohorts consisting of individuals selected for OA.

Sample Selection and Pooling Strategies – Replication Sample 1

[0217] A second case control sample (replication sample #1) was created by using 100 Caucasian female cases from Chingford, UK, and 148 unrelated female cases from the St. Thomas twin study. Cases were defined as having Kellgren-Lawrence (KL) scores of at least 2 in at least one knee x-ray. In addition, 199 male knee replacement cases from Nottingham were included. (For a cohort description, see the Nottingham description provided in Example 1). The control pool was made up of unrelated female individuals from the St. Thomas twin study (England) with normal knee x-rays and without other indications of OA, regardless of anatomical location, as well as lacking family history of OA. The St. Thomas twin study consists of Caucasian, female participants from the St. Thomas' Hospital, London, adult-twin registry, which is a voluntary registry of >4,000 twin pairs ranging from 18 to 76 years of age.

The replication sample 1 cohort was used to replicate the initial results. Table 6 below summarizes the selected phenotype data collected from the case and control individuals.

TABLE 6

Phenotype	Female cases (n=248): median (range)/ (n,%)	Male cases (n=199): median (range)/ (n,%)	Female controls (n=313): mean (range)/ (n,%)
Age	59 (39- 73)	66 (45- 73)	55 (50- 72)
Height (cm)	162 (141- 178)	175 (152- 198)	162 (141- 176)
Weight (kg)	68 (51- 123)	86 (62- 127)	64 (40- 111)
Body mass index (kg/m^2)	26 (18- 44)	29 (21- 41)	24 (18- 46)
Kellgren-Lawrence* left knee	0 (63, 26%), 1 (20, 8%), 2 (105, 43%), 3 (58, 23%), 4 (1, 0%)	NA	NA
Kellgren-Lawrence* right knee	0 (43, 7%), 1 (18, 7%), 2 (127, 52%), 3 (57, 23%), 4 (1, 0%)	NA	NA
KL* >2 both knees	No (145, 59%), Yes (101, 41%)	NA	NA
KL* >2 either knee	No (0, 0%), Yes (248, 100%)	NA	NA

* 0: normal, 1: doubtful, 2: definite osteophyte (bony protuberance), 3: joint space narrowing (with or without osteophyte), 4: joint deformity

Sample Selection and Pooling Strategies – Replication Sample 2

[0218] A third case control sample (replication sample #2) was created by using individuals with symptoms of OA from Newfoundland, Canada. These individuals were recruited and examined by rheumatologists. Affected joints were x-rayed and a final diagnosis of definite or probable OA was made according to American College of Rheumatology criteria by a single rheumatologist to avoid any inter-examiner diagnosis variability. Controls were recruited from volunteers without any symptoms from the musculoskeletal system based on a normal joint exam performed by a rheumatologist. Only cases with a diagnosis of definite OA were included in the study. Only individuals of Caucasian origin were included. The cases consisted of 228 individuals with definite knee OA, 106 individuals with definite hip OA, and 74 individuals with hip OA.

TABLE 7

Phenotype	Case	Control
Age at Visit	62.7	52.5

Phenotype	Case	Control
Sex (Female/Male)	227/119	174/101
Knee OA Xray: No	35% (120)	80% (16)
Unknown	1% (4)	0% (0)
Yes	64% (221)	20% (4)
Hip OA Xray: No	63% (215)	80% (16)
Unknown	2% (7)	0% (0)
Yes	35% (121)	20% (4)

Assay for Verifying, Allelotyping, and Genotyping SNPs

[0219] Genotyping of the replication cohorts described in Tables 6 and 7 was performed using the same methods used for the original genotyping, as described herein. A MassARRAY™ system (Sequenom, Inc.) was utilized to perform SNP genotyping in a high-throughput fashion. This genotyping platform was complemented by a homogeneous, single-tube assay method (hME™ or homogeneous MassEXTEND™ (Sequenom, Inc.)) in which two genotyping primers anneal to and amplify a genomic target surrounding a polymorphic site of interest. A third primer (the MassEXTEND™ primer), which is complementary to the amplified target up to but not including the polymorphism, was then enzymatically extended one or a few bases through the polymorphic site and then terminated.

[0220] For each polymorphism, SpectroDESIGNER™ software (Sequenom, Inc.) was used to generate a set of PCR primers and a MassEXTEND™ primer which where used to genotype the polymorphism. Other primer design software could be used or one of ordinary skill in the art could manually design primers based on his or her knowledge of the relevant factors and considerations in designing such primers. Table 3 shows PCR primers and Table 4 shows extension probes used for analyzing (e.g., genotyping) polymorphisms in the replication cohorts. The initial PCR amplification reaction was performed in a 5 µl total volume containing 1X PCR buffer with 1.5 mM MgCl₂ (Qiagen), 200 µM each of dATP, dGTP, dCTP, dTTP (Gibco-BRL), 2.5 ng of genomic DNA, 0.1 units of HotStar DNA polymerase (Qiagen), and 200 nM each of forward and reverse PCR primers specific for the polymorphic region of interest.

[0221] Samples were incubated at 95°C for 15 minutes, followed by 45 cycles of 95°C for 20 seconds, 56°C for 30 seconds, and 72°C for 1 minute, finishing with a 3 minute final extension at 72°C. Following amplification, shrimp alkaline phosphatase (SAP) (0.3 units in a 2 µl volume) (Amersham Pharmacia) was added to each reaction (total reaction volume was 7 µl) to remove any residual dNTPs

that were not consumed in the PCR step. Samples were incubated for 20 minutes at 37°C, followed by 5 minutes at 85°C to denature the SAP.

[0222] Once the SAP reaction was complete, a primer extension reaction was initiated by adding a polymorphism-specific MassEXTEND™ primer cocktail to each sample. Each MassEXTEND™ cocktail included a specific combination of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) used to distinguish polymorphic alleles from one another. Methods for verifying, allelotyping and genotyping SNPs are disclosed, for example, in U.S. Pat. No. 6,258,538, the content of which is hereby incorporated by reference. In Table 7, ddNTPs are shown and the fourth nucleotide not shown is the dNTP.

[0223] The MassEXTEND™ reaction was performed in a total volume of 9 μl, with the addition of 1X ThermoSequenase buffer, 0.576 units of ThermoSequenase (Amersham Pharmacia), 600 nM MassEXTEND™ primer, 2 mM of ddATP and/or ddCTP and/or ddGTP and/or ddTTP, and 2 mM of dATP or dCTP or dGTP or dTTP. The deoxy nucleotide (dNTP) used in the assay normally was complementary to the nucleotide at the polymorphic site in the amplicon. Samples were incubated at 94°C for 2 minutes, followed by 55 cycles of 5 seconds at 94°C, 5 seconds at 52°C, and 5 seconds at 72°C.

[0224] Following incubation, samples were desalted by adding 16 μl of water (total reaction volume was 25 μl), 3 mg of SpectroCLEAN™ sample cleaning beads (Sequenom, Inc.) and allowed to incubate for 3 minutes with rotation. Samples were then robotically dispensed using a piezoelectric dispensing device (SpectroJET™ (Sequenom, Inc.)) onto either 96-spot or 384-spot silicon chips containing a matrix that crystallized each sample (SpectroCHIP™ (Sequenom, Inc.)). Subsequently, MALDI-TOF mass spectrometry (Biflex and Autoflex MALDI-TOF mass spectrometers (Bruker Daltonics) can be used) and SpectroTYPER RT™ software (Sequenom, Inc.) were used to analyze and interpret the SNP genotype for each sample.

Genetic Analysis

[0225] Genotyping results for replication cohorts #1 and #2 are provided in Tables 8 and 9, respectively.

TABLE 8

rsID	Replication #1 (Mixed Male/Female cases and Female controls)				Meta-analysis Disc. + Rep #1
	AF OA Con	AF OA Cas	Delta	P-value	P-value
rs1041973	0.77	0.79	-0.02	0.357	NA

TABLE 9

rsID	Replication #2 (Newfoundland) (Male/Female cases and controls)				Meta-analysis Disc. + Rep #2
	AF OA Con	AF OA Cas	Delta	P-value	Not Done
rs1041973	0.78	0.79	-0.016	0.510	

[0226] To combine the evidence for association from multiple sample collections, a meta-analysis procedure was employed. The allele frequencies were compared between cases and controls within the discovery sample, as well as within the replication cohort #1 using the DerSimian-Laird approach (DerSimonian, R. and N. Laird. 1986. Meta-analysis in clinical trials. Control Clin Trials 7: 177-188.)

[0227] The absence of a statistically significant association in one or more of the replication cohorts should not be interpreted as minimizing the value of the original finding. There are many reasons why a biologically derived association identified in a sample from one population would not replicate in a sample from another population. The most important reason is differences in population history. Due to bottlenecks and founder effects, there may be common disease predisposing alleles present in one population that are relatively rare in another, leading to a lack of association in the candidate region. Also, because common diseases such as arthritis-related disorders are the result of susceptibilities in many genes and many environmental risk factors, differences in population-specific genetic and environmental backgrounds could mask the effects of a biologically relevant allele. For these and other reasons, statistically strong results in the original, discovery sample that did not replicate in one or more of the replication samples may be further evaluated in additional replication cohorts and experimental systems.

Example 4
IL1RL1 Region Proximal SNPs

[0228] It has been discovered that SNP rs1041973 in Interleukin 1 receptor-like 1 isoform 1 (*IL1RL1*) is associated with occurrence of osteoarthritis in subjects. Interleukin 1 receptor-like 1 isoform 1 is a member of the interleukin 1 receptor family with no known ligand (orphan receptor). IL1RL1 exists in both a soluble and transmembrane form, suggesting that it may have ligand or ligand scavenging activity. Studies of the similar gene in mouse suggested that this receptor can be induced by proinflammatory stimuli. This gene and four other interleukin 1 receptor family genes, including interleukin 1 receptor, type I (IL1R1), interleukin 1 receptor, type II (IL1R2), interleukin 1 receptor-like 2 (IL1RL2), and interleukin 18 receptor 1 (IL18R1), form a cytokine receptor gene cluster.

[0229] Ninety-one additional allelic variants proximal to rs1041973 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2.

The polymorphic variants are set forth in Table 10. The chromosome positions provided in column four of Table 10 are based on Genome "Build 34" of NCBI's GenBank.

TABLE 10

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs884517	2	207	102527857	c/t
rs1476984	2	6019	102533669	a/g
rs951774	2	6414	102534064	a/c
rs2041737	2	7341	102534991	a/g
rs1420091	2	10984	102538634	a/g
rs2110660	2	12351	102540001	c/g
rs1362347	2	13335	102540985	a/g
rs3073968	2	16584	102544234	-/tgtg/tgtgag
rs4090473	2	16737	102544387	c/g
rs1558622	2	23897	102551547	c/t
rs1558621	2	24057	102551707	c/t
rs1558620	2	25145	102552795	a/g
rs1558619	2	25300	102552950	a/c
rs950881	2	26262	102553912	a/c
rs950880	2	26312	102553962	g/t
rs1362346	2	26589	102554239	c/t
rs1968171	2	27302	102554952	a/g
rs1813299	2	27358	102555008	a/t
rs1813298	2	27451	102555101	c/g
rs1968170	2	27552	102555202	c/t
rs974389	2	30731	102558381	c/t
rs971764	2	32085	102559735	a/g
rs1420089	2	32139	102559789	a/g
rs1420088	2	33184	102560834	a/g
rs1420103	2	42382	102570032	g/t
rs1420102	2	42569	102570219	a/g
rs1997467	2	44823	102572473	c/t
rs1997466	2	45217	102572867	c/g
rs1362350	2	45548	102573198	c/g
rs2310220	2	45601	102573251	a/g
rs1362349	2	45722	102573372	c/g
rs3755278	2	45967	102573617	a/g
rs3771180	2	47367	102575017	a/c
rs3771179	2	47642	102575292	a/c
rs985523	2	48126	102575776	c/t
rs1041973	2	49218	102576868	a/c
rs3214363	2	49274	102576924	-/a
rs873022	2	49433	102577083	g/t
rs3771177	2	49610	102577260	a/c
rs3732129	2	51282	102578932	a/g
rs1420101	2	51466	102579116	a/g
rs12905	2	53757	102581407	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs3771175	2	53960	102581610	a/t
rs3821204	2	54031	102581681	c/g
rs2160203	2	54574	102582224	c/t
rs1946131	2	55679	102583329	a/g
rs1054096	2	56100	102583750	c/t
rs2287038	2	56182	102583832	c/t
rs1921622	2	59817	102587467	a/g
rs1861246	2	60533	102588183	a/g
rs1861245	2	60656	102588306	a/g
rs3755276	2	72209	102599859	a/g
rs2287037	2	72778	102600428	a/g
rs1420099	2	74293	102601943	c/g
rs3771174	2	77335	102604985	a/g
rs1420098	2	78029	102605679	a/g
rs1362348	2	78374	102606024	c/g
rs1882348	2	78421	102606071	a/t
rs1558627	2	78434	102606084	c/t
rs2058622	2	79174	102606824	c/t
rs3836110	2	79397	102607047	-/g
rs3771172	2	79562	102607212	a/g
rs3771171	2	79700	102607350	a/g
rs3771170	2	79730	102607380	a/t
rs2160202	2	79904	102607554	c/t
rs2058623	2	79920	102607570	a/g
rs3771167	2	79938	102607588	c/t
rs3771166	2	79972	102607622	c/t
rs1974675	2	80125	102607775	c/t
rs1465321	2	80368	102608018	a/g
rs2041740	2	83484	102611134	c/t
rs3771164	2	85536	102613186	a/t
rs2270298	2	85829	102613479	c/t
rs2270297	2	86425	102614075	a/g
rs2041739	2	88083	102615733	a/g
rs2080289	2	88770	102616420	c/t
rs3821203	2	90622	102618272	a/g
rs3771162	2	90924	102618574	a/t
rs3213733	2	91634	102619284	g/t
rs3213732	2	92029	102619679	c/t
rs1035130	2	95152	102622802	a/g
rs3752659	2	95348	102622998	c/t
rs3755274	2	96145	102623795	c/t
rs2241117	2	96793	102624443	a/g
rs2241116	2	97015	102624665	g/t
rs881890	2	97064	102624714	c/t
rs3771161	2	97711	102625361	g/t
rs3771160	2	97855	102625505	a/c
rs3771159	2	98708	102626358	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs1420104	2	not mapped	not mapped	c/t
rs2041738	2	not mapped	not mapped	a/c

Assay for Verifying and Allelotyping SNPs

[0230] The methods used to verify and allelotype the 91 proximal SNPs of Table 10 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 11 and Table 12, respectively.

TABLE 11

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs884517	ACGTTGGATGCATTTTCTGGTGTGACTCCC	ACGTTGGATGATGTTCGGTCACTTGTGAGC
rs1476984	ACGTTGGATGTGAGAGAGTTGAAGAATGGG	ACGTTGGATGCCAAGAAGTGATTTCCTTCC
rs951774	ACGTTGGATGTCAGCCAGAGGTCTTTACTC	ACGTTGGATGTTAGAAGTCTCTTGGGTGGG
rs2041737	ACGTTGGATGGAGATGGAGTTTCCCTCTTG	ACGTTGGATGAAACCAAGAGGTGGAGGTTG
rs1420091	ACGTTGGATGCACCCCTATTATAAAACCCAC	ACGTTGGATGACCAGAAATGGCATCTATGG
rs2110660	ACGTTGGATGTCTCTCCGAGATGAGGAATC	ACGTTGGATGGTGATCTCCTCAGTACTCTG
rs1362347	ACGTTGGATGTTCTTTGGTAATGAGGTAGG	ACGTTGGATGTGCTTGCCCTCTATTTATGG
rs3073968	ACGTTGGATGGAATGATGAGGAAGGAAGGG	ACGTTGGATGTAAAGCCACATGTTCACCCG
rs4090473	ACGTTGGATGTAGTGTGTTTCACTCTTCCC	ACGTTGGATGTCAAGCACCTCTGTTAACTC
rs1558622	ACGTTGGATGATACTTCCTGGTTTTCTGGG	ACGTTGGATGGGCTCAAAGTCATCACCCAA
rs1558621	ACGTTGGATGACAGTGGCGATGCCAACATT	ACGTTGGATGCCTGTAGTAGGACCCTACTG
rs1558620	ACGTTGGATGTTGCAGGTGTCTGGTGATAG	ACGTTGGATGAGTTTGCCTTTCTTCATGGC
rs1558619	ACGTTGGATGCCCTAATTAGGATTCCGCAC	ACGTTGGATGCTCCATCACACTTTGACTGC
rs950881	ACGTTGGATGCTTATCTCAGTCTGCCAGTG	ACGTTGGATGGGTGAGTGAATTAGTCCTGG
rs950880	ACGTTGGATGTGCCAAAGACAATCAAATCC	ACGTTGGATGCACTCACCTCTGATTTCTAG
rs1362346	ACGTTGGATGTTCCTCAGGTTCACCAAGAG	ACGTTGGATGTCCCGAACCTCATCTCATAC
rs1968171	ACGTTGGATGAATGTTTCAGCCCAGCATGG	ACGTTGGATGATCTCCTGACCTCATGATCC
rs1813299	ACGTTGGATGAATTCCAGCACTTTGGGAGG	ACGTTGGATGTTTCACCGTGTTAGCCAGGA
rs1813298	ACGTTGGATGTTACTGCAAGCTCCACCTCC	ACGTTGGATGATTAACTGGGCGTGGTGGTG
rs1968170	ACGTTGGATGAGCTTGCAGTAAGCCCAGAT	ACGTTGGATGTGTTAGGGTAATTACAGTGC
rs974389	ACGTTGGATGCTCTAGCCCAATATGTCTCC	ACGTTGGATGACTGGAGATGTGAACCCATC
rs971764	ACGTTGGATGGAGATGATGGAGATTAAGAGG	ACGTTGGATGAGTTGTTTGACTTCGGACTG
rs1420089	ACGTTGGATGAGACAGCACATATCAATGAC	ACGTTGGATGTATTGTGCGGTTCGCTATAG
rs1420088	ACGTTGGATGGGATGACTGTCAAAAACATC	ACGTTGGATGTAATTTTTCAGGAGCAAGGC
rs1420103	ACGTTGGATGTCCATTGGAATATGACCTCC	ACGTTGGATGCCAGGCACATGAGCTATATC
rs1420102	ACGTTGGATGGATTGGTCAGGAACTCAAAC	ACGTTGGATGTGGGTTGCTTCTAGCTATTG
rs1997467	ACGTTGGATGTGAATTTCAGTGAGTCAGGC	ACGTTGGATGTGAGGGGAAAAAAAACATCC
rs1997466	ACGTTGGATGATAGGCACATACAGGATTTC	ACGTTGGATGCTCCCTTTTCAGATTAATCTC
rs1362350	ACGTTGGATGGAGAACATTCTCTATACCAG	ACGTTGGATGTGCCTGAATAGTGAGAAGCC
rs2310220	ACGTTGGATGGGTTGAAACCAGACTTGCTG	ACGTTGGATGCAGCCTAATCTCTGGTATAG
rs1362349	ACGTTGGATGCAATACTCTGTGGTACTTATC	ACGTTGGATGTAAACAGTCTTATCCTTGGG
rs3755278	ACGTTGGATGAGTGCTGAATAGGTTTGTTC	ACGTTGGATGGCCTAGTTTAAGAATGAATGC

68

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs3771180	ACGTTGGATGGTCAACATCAAGAATTCTTAG	ACGTTGGATGCCTGAAATTTGATTTGTGGC
rs3771179	ACGTTGGATGGTCTTCATAATTCATGATTG	ACGTTGGATGTCTTAAAATATAAGGGGAAG
rs985523	ACGTTGGATGTCCTATGGAAGTTTTGGGTC	ACGTTGGATGCTGCGAAGTAGCATGATAAC
rs1041973	ACGTTGGATGGGGACTTCTGACAATACAGG	ACGTTGGATGAATCGTGTGTTTGCCTCAGG
rs3214363	ACGTTGGATGCAGGCAATCAACCACTGAAG	ACGTTGGATGCTGCAGTTGCTGATTCTGGT
rs873022	ACGTTGGATGCCTAGTCCTTTCTGGAACAG	ACGTTGGATGATCCCTGCAACTGTAAATCC
rs3771177	ACGTTGGATGAAGGTTAGAAGCCCCTTTTC	ACGTTGGATGGGGCTGGAATTAAGAACAAAC
rs3732129	ACGTTGGATGCTAATTCAAAGCCACATCTG	ACGTTGGATGTAAGTTAGCATTCAGATTGC
rs1420101	ACGTTGGATGCAACATTTATGTACACCATAG	ACGTTGGATGTTAGTAATACTCATTGGATT
rs12905	ACGTTGGATGCTCCCAGCAAACAGGAACAG	ACGTTGGATGATCAAGACAATGGGAATGGC
rs3771175	ACGTTGGATGAAAGAGCACAAAGAACACG	ACGTTGGATGTTATGAACTCCCTCTGTGTC
rs3821204	ACGTTGGATGCATGTTGTAAGCATGGTCCG	ACGTTGGATGACTTTACCACCCTCGCTAAC
rs2160203	ACGTTGGATGACACAGACCCAAACCATACC	ACGTTGGATGTTCCCGTGTGTTCCATGTAC
rs1946131	ACGTTGGATGGGGAACTCAGGGTTTAACAC	ACGTTGGATGTACACTCATCACTCCTCAGG
rs1054096	ACGTTGGATGATCAAGGTGCTATGTGAGGG	ACGTTGGATGAAAGCAGGAGTACACAAGGG
rs2287038	ACGTTGGATGAATGTCCCTGGTTACCTATG	ACGTTGGATGACAAATAAGCTAGAAGGAGC
rs1921622	ACGTTGGATGGCCACTTCTTAATTCTGTCC	ACGTTGGATGATTTCAGCTAGTGCCTATGG
rs1861246	ACGTTGGATGCACAAGCTCTTCACCTCTTC	ACGTTGGATGTGGCTGAGGAGAAGTGTAAC
rs1861245	ACGTTGGATGTGCTGCCTTCAATGTGTGAC	ACGTTGGATGAGGAAAGGTCAGAGGACATG
rs3755276	ACGTTGGATGCCAGCACTCACTAACATGTG	ACGTTGGATGAAACTCATATGGGCAGCCAC
rs2287037	ACGTTGGATGCAGATTCAGCCAAAGCTTTC	ACGTTGGATGAAAAATCTGTGTGCCAGAAG
rs1420099	ACGTTGGATGTTACACACTCTCCAGAGGTG	ACGTTGGATGAAAGCTTCTAGCTGCCTGAG
rs3771174	ACGTTGGATGACCCAGATTCTCTGGCTTTG	ACGTTGGATGTACCACAAGTGCCGAAAGAG
rs1420098	ACGTTGGATGGGGACGTGAAGTACAAGAT	ACGTTGGATGGGAGACCAAAAAAAGTTACC
rs1362348	ACGTTGGATGCATGTCATAGGAAGAGTAGG	ACGTTGGATGTCAGCAACTCAAATATGCAG
rs1882348	ACGTTGGATGCCTACTCTTCCTATGACATG	ACGTTGGATGCCCTAAAAGGAAATCCTATC
rs1558627	ACGTTGGATGCCCTAAAAGGAAATCCTATC	ACGTTGGATGCCTACTCTTCCTATGACATG
rs2058622	ACGTTGGATGCTGTGAAACCTTGGTAGCAC	ACGTTGGATGTTTCTGATGCCTGGGAGTTC
rs3836110	ACGTTGGATGACTCACAAATGGGGTAAAGG	ACGTTGGATGTGCCTTCATTCAATCAGGAG
rs3771172	ACGTTGGATGCAGAAGCAAATGGCATTGGC	ACGTTGGATGCCATTGTTGCTTCCTAAGCC
rs3771171	ACGTTGGATGAGGGTAGCAGATAGGAGATG	ACGTTGGATGAAGCTGCTTCTCTCCTCATC
rs3771170	ACGTTGGATGCAAGGCCATTGTCAAAGCTG	ACGTTGGATGGGTGTCCCAGAGTGGATATTG
rs2160202	ACGTTGGATGAGCAGTATTTACTGCAGATG	ACGTTGGATGCCCACATCAAACTGCAAAGG
rs2058623	ACGTTGGATGATTTACTGCAGATGTGTGTG	ACGTTGGATGTGTTCACTGATAGATCCCAC
rs3771167	ACGTTGGATGCTAACTTAAGTGTGTAACCC	ACGTTGGATGCTAACGGGAAATTTTCAGGTG
rs3771166	ACGTTGGATGGTGAACAGACTTTACACCTG	ACGTTGGATGCCTCAGTGGCATTTGATTAT
rs1974675	ACGTTGGATGACTAAGAAGGAAGGGGATAC	ACGTTGGATGGTACATTTCCCTCTACCTTC
rs1465321	ACGTTGGATGTCACAGCTTTGGGTCAGTTG	ACGTTGGATGTCAACAACACACTGCACCTG
rs2041740	ACGTTGGATGCATCCATGTCCCTACAAAAG	ACGTTGGATGAAAGCTCTTATACACCATGG
rs3771164	ACGTTGGATGCCTGTGACATGTATGGAAATG	ACGTTGGATGTCAAATCCATAGGTACACTC
rs2270298	ACGTTGGATGTGAAGTAGTGTTCTCTCTC	ACGTTGGATGAATATGAGCACTGTAGCTGC
rs2270297	ACGTTGGATGTTTCCTGCCAAAAAGAAAGG	ACGTTGGATGGACCACACCACTAGTTCAAA
rs2041739	ACGTTGGATGTAGACCCTGAAGTTTCCCAC	ACGTTGGATGCACCTAGAGGTTCCTTTTGC
rs2080289	ACGTTGGATGTGGAGAATGTCAACTGAGTC	ACGTTGGATGATACAAACAAGAGGCCATGG
rs3821203	ACGTTGGATGTCAAAGACAAAGGGCAGGAG	ACGTTGGATGGGATCCAGAGAAAGGTAGTC
rs3771162	ACGTTGGATGTGAGTGGAGTACAGTGAGAC	ACGTTGGATGTGGCACTGCACTTTCTGAGA
rs3213733	ACGTTGGATGTGAAAGCACCTTGTATCTGG	ACGTTGGATGCATCTTCCTCTGCCTTTTAG

69

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs3213732	ACGTTGGATGGTCAGGTTAAAAGTGGCAAC	ACGTTGGATGTGACACTGGATACACATTTC
rs1035130	ACGTTGGATGTTAGGATCCGATCCATTTTC	ACGTTGGATGCTCTGCTTTGCTGAATGAAG
rs3752659	ACGTTGGATGTGCATAATGCGTCCACCTAG	ACGTTGGATGGGCTGATGTGTATTTTGGGC
rs3755274	ACGTTGGATGTATCAAAGGTGTGTGCACCC	ACGTTGGATGAGGGGTAGAAAACCACAGTG
rs2241117	ACGTTGGATGTGGCTGGAAGATCATGATGC	ACGTTGGATGCCCCAAGTTGTTAGGAAGAG
rs2241116	ACGTTGGATGAATGCAGGCAACATCACAGC	ACGTTGGATGAGTAGGCTCTGTTCGTTACC
rs881890	ACGTTGGATGATGCCATTTGCCTTCTGGAG	ACGTTGGATGTCTCAGGGTAACGAACAGAG
rs3771161	ACGTTGGATGCCATCAGGTGAGCACTGAAA	ACGTTGGATGTCATTGCCTCCTGAACTTGG
rs3771160	ACGTTGGATGAGAAATGGCTGTGACTGGAG	ACGTTGGATGTATCCAGGGAGTTGATGGTG
rs3771159	ACGTTGGATGCAGGTGATGGTCCAACAAAG	ACGTTGGATGTGCTGTGGTCCACTCACTTG
rs1420104	ACGTTGGATGTATTCTGGAGGCTGAGGTGG	ACGTTGGATGTGGAGTGCAGTGGTGTGATC
rs2041738	ACGTTGGATGTGGTGAAACCCCATCTCTAC	ACGTTGGATGTTTCAAGCTATTCTCCTGCC

TABLE 12

dbSNP rs#	Extend Primer	Term Mix
rs884517	GGTGTGACTCCCAGACCAA	ACT
rs1476984	ATGGGTAGTTAATGGTGGAAATTT	ACT
rs951774	CAAAGTAGTTGACTTGTCTTTCT	ACT
rs2041737	CCAGGCTAGTGCAGTGGC	ACT
rs1420091	CCCACATTATATTGTCATTACTTT	ACG
rs2110660	ATGAGGAATCAGAGCTGGGA	ACT
rs1362347	GTAATGAGGTAGGAATAATATTG	ACT
rs3073968	GGCAATTGTGTGTGTGTGTG	CGT
rs4090473	CTTACTCCTATTCCAAAGTTCA	ACT
rs1558622	ACTGCAAGGGAGAGCCCC	ACT
rs1558621	AGTGTGTGTGTGTGCGTGC	ACT
rs1558620	GTCTGGTGATAGTTGGGTGC	ACG
rs1558619	GATTCCGCACATCCTATGCCT	ACT
rs950881	GATGGTTTGTGCCTCTGGTC	ACT
rs950880	ATTTAAGAATGCTTTCGTCATAAG	ACT
rs1362346	GAATATCTATGCCCACCAGAT	ACG
rs1968171	GCCCAGCATGGTGGCTCA	ACG
rs1813299	GTGGATCATGAGGTCAGGAG	CGT
rs1813298	GCCTCAGCCTCCCGAGTA	ACT
rs1968170	AGCCTGGGTGACAGAGCC	ACT
rs974389	GTCTCCTGAATTTCAGAAGCA	ACT
rs971764	GTCAAGGTAAAAACATTATTGTG	ACG
rs1420089	GCACATATCAATGACAAGACTA	ACT
rs1420088	CATGTTATGTAACTCTGAGTTC	ACT
rs1420103	GAATATGACCTCCAGAAGGCAA	ACT
rs1420102	GAACTCAAACAAATACTTGGACAC	ACG
rs1997467	TTCAGTGACTCTCACAATAAGC	ACG
rs1997466	AAGAAAAAGCTGGTTCAATGAG	ACT
rs1362350	ACATTCTCTATACCAGAGATTAGG	ACT

70

dbSNP rs#	Extend Primer	Term Mix
rs2310220	CTGAACTTCAAAGTCAAGCTTTT	ACG
rs1362349	CTGTGGTACTTATCATTAACATCA	ACT
rs3755278	ACTCGGAATTCTTTTACATTTGGT	ACT
rs3771180	CATCAAGAATTCTTAGTACATGAT	ACT
rs3771179	TATGTTAGTAAATTTCTATGTTGG	ACT
rs985523	CATATAGCTTTCACAATGATCATG	ACG
rs1041973	ATACCAGAATCAGCAACT	ACT
rs3214363	GAGCAGGGTGAAAGAAGATGGG	ACT
rs873022	TTCTAGGAATACTATCAGGTTGA	ACT
rs3771177	TTTTCACCTACTAGAGGCCC	CGT
rs3732129	GCCACATCTGTTCTTTATTCTTT	ACG
rs1420101	CCATCACAAAGCCTCTCATTA	ACT
rs12905	AGACAGCAAACAACATCC	ACG
rs3771175	CACAAAAGAACACGTTCAGTTT	CGT
rs3821204	TAAGCATGGTCCGTTCTATAC	ACT
rs2160203	CCACACACATTATCATTGTTA	ACT
rs1946131	TTAACACTCTTTGGCTATTTGACA	ACT
rs1054096	TCCATCCAGCCTGCCCAC	ACG
rs2287038	TACCTATGTGTTTGAATTATCTTC	ACT
rs1921622	GAAAGAGGACTTAAAAATTGATGA	ACT
rs1861246	CTTCACCTCTTCTTTTTCAGTC	ACG
rs1861245	CTGGAATGGTTTTCTACTTCC	ACG
rs3755276	GTGTGTATGCATGTGTTCGC	ACT
rs2287037	ACAAAAGTGTGCCTATCTTATGAA	ACT
rs1420099	GGTGGGAGGTTGATAATTGAAA	ACT
rs3771174	CTGACCATCATCTACCCAGG	ACT
rs1420098	ACGTGAAGTACAAGATTCTTCA	ACT
rs1362348	GAGTAGGAAAGAAAAGGATGTG	ACT
rs1882348	TCCTATGACATGAAATACATTCT	CGT
rs1558627	AAGCAGAGAGAGATAAACTTATT	ACG
rs2058622	AAACCTTGGTAGCACTTCTGT	ACT
rs3836110	AACAAACACCGCCCCCCC	CGT
rs3771172	GCATTGGCCATCTTTCTGATA	ACG
rs3771171	GAGGTGTCCCAGAGTGGATA	ACG
rs3771170	CAAAGCTGCTTCTCTCCTCA	CGT
rs2160202	TATACACATATGTGTTCTAACTTA	ACT
rs2058623	ACTTAGGTGTGTAACCCTTTG	ACG
rs3771167	CTTTGTAGTTTGATGTGGGATCT	ACT
rs3771166	ACTTTACACCTGAAAATTTCCC	ACT
rs1974675	GAAGGGGATACAAAAGGGATA	ACT
rs1465321	CAGTTGGCCTCAGTGTTAACCC	ACG
rs2041740	GAACTCATGCTTTTTTATGGCTG	ACG
rs3771164	GACATGTATGGAAATGTGTGTG	CGT
rs2270298	CTCTCTCTCTGCATGTGTGT	ACT

dbSNP rs#	Extend Primer	Term Mix
rs2270297	AGCCAAGTAGAGGAGCACC	ACT
rs2041739	CTCCTGAGTTCCTGTGAATAC	ACT
rs2080289	TCTCAGGACTCCACTCAAATGTC	ACT
rs3821203	GGCAGGAGGCAATTTCGGT	ACT
rs3771162	CAGTGAGACTCAGGAGTGC	CGT
rs3213733	TGTATCTGGTTTTCTCTCACTCA	ACT
rs3213732	CAACATTCAAAAAATGGCACTCTT	ACG
rs1035130	TCCGATCCATTTTCTTCCCC	ACT
rs3752659	CCTAGGGTATGGCCACTATAATTA	ACG
rs3755274	CACCCAACTATAAAGAAAGACCTC	ACG
rs2241117	ATCATGATGCTAAGTTGAAAATAT	ACT
rs2241116	TCAAGCATTTTAAACATGTGAATT	CGT
rs881890	TGCCTTCTGGAGTCCTGTAA	ACT
rs3771161	GTGAGCACTGAAAAACTTTAAGA	ACT
rs3771160	GCCAGAAAGCTGTGATTTCCA	ACT
rs3771159	CCAACAAAGATTTGAGCCCC	ACT
rs1420104	CTGGGAGGTGGAGACTGCA	ACT
rs2041738	AAAAATACAAAAATTAGCTGGGC	ACT

Genetic Analysis

[0231] Allelotyping results from the discovery cohort are shown for cases and controls in Table 13. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where "AF" is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs951774 has the following case and control allele frequencies: case A1 (A) = 0.24; case A2 (C) = 0.76; control A1 (A) = 0.20; and control A2 (C) = 0.80, where the nucleotide is provided in paranthesis. Some SNPs are labeled "untyped" because of failed assays.

TABLE 13

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p-Value
rs884517	207	102527857	C/T			
rs1476984	6019	102533669	A/G	0.83	0.83	0.973
rs951774	6414	102534064	A/C	0.76	0.80	0.099
rs2041737	7341	102534991	A/G	0.38	0.32	0.146
rs1420091	10984	102538634	A/G	0.33	0.35	0.388
rs2110660	12351	102540001	C/G	0.41	0.40	0.753
rs1362347	13335	102540985	A/G	0.83	0.83	0.895
rs3073968	16584	102544234	-/TGTG/T GTGAG	0.48	0.48	0.878
rs4090473	16737	102544387	C/G	0.42	0.43	0.633
rs1558622	23897	102551547	C/T	0.40	0.39	0.879

72

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p-Value
rs1558621	24057	102551707	C/T	0.32	0.31	0.795
rs1558620	25145	102552795	A/G	0.37	0.37	0.998
rs1558619	25300	102552950	A/C	0.46	0.47	0.556
rs950881	26262	102553912	A/C	0.75	0.74	0.636
rs950880	26312	102553962	G/T	0.45	0.48	0.285
rs1362346	26589	102554239	C/T			
rs1968171	27302	102554952	A/G	0.43	0.43	0.891
rs1813299	27358	102555008	A/T			
rs1813298	27451	102555101	C/G			
rs1968170	27552	102555202	C/T	0.65	0.65	0.941
rs974389	30731	102558381	C/T	0.41	0.42	0.734
rs971764	32085	102559735	A/G	0.45	0.44	0.738
rs1420089	32139	102559789	A/G	0.16	0.19	0.099
rs1420088	33184	102560834	A/G	0.41	0.40	0.869
rs1420103	42382	102570032	G/T	0.68	0.68	0.952
rs1420102	42569	102570219	A/G	0.48	0.46	0.349
rs1997467	44823	102572473	C/T			
rs1997466	45217	102572867	C/G	0.46	0.46	0.693
rs1362350	45548	102573198	C/G	0.48	0.46	0.475
rs2310220	45601	102573251	A/G	0.40	0.41	0.480
rs1362349	45722	102573372	C/G	0.41	0.42	0.893
rs3755278	45967	102573617	A/G	0.07	0.08	0.876
rs3771180	47367	102575017	A/C	0.91	0.90	0.669
rs3771179	47642	102575292	A/C	0.08	0.08	0.986
rs985523	48126	102575776	C/T	0.17	0.13	0.064
rs1041973	49218	102576868	A/C			
rs3214363	49274	102576924	-/A			
rs873022	49433	102577083	G/T	0.53	0.56	0.321
rs3771177	49610	102577260	A/C	0.33	0.31	0.278
rs3732129	51282	102578932	A/G	0.46	0.50	0.127
rs1420101	51466	102579116	A/G	0.55	0.57	0.257
rs12905	53757	102581407	A/G	0.30	0.27	0.262
rs3771175	53960	102581610	A/T	0.84	0.82	0.174
rs3821204	54031	102581681	C/G	0.26	0.23	0.222
rs2160203	54574	102582224	C/T	0.21	0.26	**0.033**
rs1946131	55679	102583329	A/G	0.73	0.74	0.710
rs1054096	56100	102583750	C/T	0.69	0.65	0.137
rs2287038	56182	102583832	C/T	0.98	0.95	0.207
rs1921622	59817	102587467	A/G	0.40	0.43	0.218
rs1861246	60533	102588183	A/G	0.22	0.18	0.068
rs1861245	60656	102588306	A/G	0.35	0.37	0.377
rs3755276	72209	102599859	A/G	0.51	0.48	0.355
rs2287037	72778	102600428	A/G	0.49	0.53	0.195
rs1420099	74293	102601943	C/G	0.58	0.56	0.416
rs3771174	77335	102604985	A/G			
rs1420098	78029	102605679	A/G	0.33	0.32	0.532
rs1362348	78374	102606024	C/G	0.02	0.03	0.590
rs1882348	78421	102606071	A/T	0.36	0.35	0.596
rs1558627	78434	102606084	C/T	0.62	0.65	0.219
rs2058622	79174	102606824	C/T	0.57	0.59	0.528
rs3836110	79397	102607047	-/G	0.72	0.73	0.856
rs3771172	79562	102607212	A/G	0.28	0.25	0.261
rs3771171	79700	102607350	A/G			
rs3771170	79730	102607380	A/T	0.24	0.23	0.533
rs2160202	79904	102607554	C/T	0.55	0.62	0.061
rs2058623	79920	102607570	A/G	0.67	0.68	0.631
rs3771167	79938	102607588	C/T			

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p-Value
rs3771166	79972	102607622	C/T	0.55	0.53	0.624
rs1974675	80125	102607775	C/T	0.57	0.55	0.470
rs1465321	80368	102608018	A/G	0.27	0.26	0.614
rs2041740	83484	102611134	C/T	0.26	0.25	0.622
rs3771164	85536	102613186	A/T	0.76	0.73	0.197
rs2270298	85829	102613479	C/T	0.23	0.21	0.329
rs2270297	86425	102614075	A/G	0.60	0.60	0.900
rs2041739	88083	102615733	A/G	0.43	0.40	0.235
rs2080289	88770	102616420	C/T	0.56	0.59	0.322
rs3821203	90622	102618272	A/G	0.58	0.62	0.194
rs3771162	90924	102618574	A/T	0.30	0.28	0.260
rs3213733	91634	102619284	G/T	0.76	0.73	0.287
rs3213732	92029	102619679	C/T	0.44	0.42	0.507
rs1035130	95152	102622802	A/G	0.58	0.61	0.234
rs3752659	95348	102622998	C/T	0.80	0.80	0.957
rs3755274	96145	102623795	C/T	0.26	0.25	0.549
rs2241117	96793	102624443	A/G	0.71	0.75	0.077
rs2241116	97015	102624665	G/T	0.16	0.15	0.469
rs881890	97064	102624714	C/T			
rs3771161	97711	102625361	G/T	0.70	0.68	0.348
rs3771160	97855	102625505	A/C			
rs3771159	98708	102626358	A/G	0.38	0.40	0.294
rs1420104	not mapped	not mapped	C/T			
rs2041738	not mapped	not mapped	A/C			

[0232] The *IL1RL1* proximal SNPs were also allelotyped in the replication cohorts using the methods described herein and the primers provided in Tables 11 and 12. The replication allelotyping results for replication cohort #1 and replication cohort #2 are provided in Tables 14 and 15, respectively.

TABLE 14

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p-Value
rs884517	207	102527857	C/T			
rs1476984	6019	102533669	A/G	0.82	0.81	0.878
rs951774	6414	102534064	A/C	0.76	0.82	0.024
rs2041737	7341	102534991	A/G	0.36	0.32	0.382
rs1420091	10984	102538634	A/G	0.32	0.35	0.340
rs2110660	12351	102540001	C/G	0.39	0.39	0.951
rs1362347	13335	102540985	A/G	0.83	0.83	0.766
rs3073968	16584	102544234	-/TGTG/T GTGAG	0.47	0.48	0.822
rs4090473	16737	102544387	C/G	0.40	0.41	0.663
rs1558622	23897	102551547	C/T	0.38	0.38	0.943
rs1558621	24057	102551707	C/T	0.33	0.32	0.631
rs1558620	25145	102552795	A/G	0.34	0.34	0.957
rs1558619	25300	102552950	A/C	0.44	0.47	0.368
rs950881	26262	102553912	A/C	0.76	0.74	0.476
rs950880	26312	102553962	G/T	0.42	0.47	0.199
rs1362346	26589	102554239	C/T			
rs1968171	27302	102554952	A/G	0.43	0.44	0.641
rs1813299	27358	102555008	A/T			

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p-Value
rs1813298	27451	102555101	C/G			
rs1968170	27552	102555202	C/T	0.64	0.65	0.797
rs974389	30731	102558381	C/T	0.39	0.41	0.678
rs971764	32085	102559735	A/G	0.47	0.46	0.710
rs1420089	32139	102559789	A/G	0.16	0.21	0.075
rs1420088	33184	102560834	A/G	0.41	0.40	0.869
rs1420103	42382	102570032	G/T	0.69	0.72	0.268
rs1420102	42569	102570219	A/G	0.50	0.47	0.329
rs1997467	44823	102572473	C/T			
rs1997466	45217	102572867	C/G	0.49	0.47	0.675
rs1362350	45548	102573198	C/G	0.51	0.47	0.308
rs2310220	45601	102573251	A/G	0.40	0.44	0.282
rs1362349	45722	102573372	C/G	0.42	0.43	0.730
rs3755278	45967	102573617	A/G	0.08	0.08	0.902
rs3771180	47367	102575017	A/C	0.93	0.92	0.591
rs3771179	47642	102575292	A/C	0.08	0.08	0.936
rs985523	48126	102575776	C/T	0.17	0.13	0.156
rs1041973	49218	102576868	A/C			
rs3214363	49274	102576924	-/A			
rs873022	49433	102577083	G/T	0.51	0.56	0.138
rs3771177	49610	102577260	A/C	0.36	0.32	0.125
rs3732129	51282	102578932	A/G	0.43	0.50	**0.048**
rs1420101	51466	102579116	A/G	0.50	0.55	0.132
rs12905	53757	102581407	A/G	0.33	0.28	0.127
rs3771175	53960	102581610	A/T	0.86	0.83	0.217
rs3821204	54031	102581681	C/G	0.29	0.23	0.071
rs2160203	54574	102582224	C/T	0.19	0.26	**0.016**
rs1946131	55679	102583329	A/G	0.72	0.73	0.771
rs1054096	56100	102583750	C/T	0.70	0.65	0.079
rs2287038	56182	102583832	C/T	0.93	NA	0.975
rs1921622	59817	102587467	A/G	0.37	0.41	0.260
rs1861246	60533	102588183	A/G	0.22	0.15	0.031
rs1861245	60656	102588306	A/G	0.34	0.39	0.149
rs3755276	72209	102599859	A/G	0.53	0.46	0.072
rs2287037	72778	102600428	A/G	0.45	0.51	0.069
rs1420099	74293	102601943	C/G	0.59	0.55	0.312
rs3771174	77335	102604985	A/G			
rs1420098	78029	102605679	A/G	0.35	0.32	0.328
rs1362348	78374	102606024	C/G	0.02	NA	**0.025**
rs1882348	78421	102606071	A/T	0.40	0.37	0.399
rs1558627	78434	102606084	C/T	0.64	0.69	0.118
rs2058622	79174	102606824	C/T	0.59	0.62	0.491
rs3836110	79397	102607047	-/G	0.74	0.75	0.625
rs3771172	79562	102607212	A/G	0.31	0.27	0.200
rs3771171	79700	102607350	A/G			
rs3771170	79730	102607380	A/T	0.22	0.20	0.346
rs2160202	79904	102607554	C/T	0.55	0.60	0.217
rs2058623	79920	102607570	A/G	0.69	0.72	0.266
rs3771167	79938	102607588	C/T			
rs3771166	79972	102607622	C/T	0.57	untyped	NA
rs1974675	80125	102607775	C/T	0.58	0.54	0.297
rs1465321	80368	102608018	A/G	0.25	0.23	0.471
rs2041740	83484	102611134	C/T	0.25	0.22	0.450
rs3771164	85536	102613186	A/T	0.77	0.72	0.073
rs2270298	85829	102613479	C/T	0.25	0.22	0.324
rs2270297	86425	102614075	A/G	0.63	0.64	0.589
rs2041739	88083	102615733	A/G	0.44	0.40	0.157

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p-Value
rs2080289	88770	102616420	C/T	0.53	0.58	0.114
rs3821203	90622	102618272	A/G	0.55	0.61	0.104
rs3771162	90924	102618574	A/T	0.34	0.29	0.261
rs3213733	91634	102619284	G/T	0.77	0.73	0.260
rs3213732	92029	102619679	C/T	0.48	0.41	**0.026**
rs1035130	95152	102622802	A/G	0.55	0.60	0.152
rs3752659	95348	102622998	C/T	0.81	0.80	0.760
rs3755274	96145	102623795	C/T	0.25	0.22	0.319
rs2241117	96793	102624443	A/G	0.72	0.80	**0.024**
rs2241116	97015	102624665	G/T	0.18	NA	NA
rs881890	97064	102624714	C/T			
rs3771161	97711	102625361	G/T	0.71	0.66	0.146
rs3771160	97855	102625505	A/C			
rs3771159	98708	102626358	A/G	0.38	0.42	0.175
rs1420104	not mapped	not mapped	C/T			
rs2041738	not mapped	not mapped	A/C			

TABLE 15

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p-Value
rs884517	207	102527857	C/T			
rs1476984	6019	102533669	A/G	0.84	0.85	0.759
rs951774	6414	102534064	A/C	0.77	0.75	0.730
rs2041737	7341	102534991	A/G	0.40	NA	
rs1420091	10984	102538634	A/G	0.34	0.35	0.819
rs2110660	12351	102540001	C/G	0.43	0.42	0.665
rs1362347	13335	102540985	A/G	0.82	0.83	0.576
rs3073968	16584	102544234	-/TGTG/T GTGAG	0.49	0.49	0.997
rs4090473	16737	102544387	C/G	0.44	0.45	0.732
rs1558622	23897	102551547	C/T	0.42	0.42	0.976
rs1558621	24057	102551707	C/T	0.31	0.31	0.926
rs1558620	25145	102552795	A/G	0.42	0.42	0.867
rs1558619	25300	102552950	A/C	0.48	0.48	0.930
rs950881	26262	102553912	A/C	0.73	0.73	0.955
rs950880	26312	102553962	G/T	0.48	0.49	0.919
rs1362346	26589	102554239	C/T			
rs1968171	27302	102554952	A/G	0.43	0.42	0.717
rs1813299	27358	102555008	A/T			
rs1813298	27451	102555101	C/G			
rs1968170	27552	102555202	C/T	0.67	0.66	0.830
rs974389	30731	102558381	C/T	0.44	0.45	0.857
rs971764	32085	102559735	A/G	0.43	0.42	0.845
rs1420089	32139	102559789	A/G	0.15	0.16	0.809
rs1420088	33184	102560834	A/G			
rs1420103	42382	102570032	G/T	0.68	0.63	0.178
rs1420102	42569	102570219	A/G	0.45	0.44	0.722
rs1997467	44823	102572473	C/T			
rs1997466	45217	102572867	C/G	0.44	0.43	0.805
rs1362350	45548	102573198	C/G	0.45	0.46	0.890
rs2310220	45601	102573251	A/G	0.38	0.37	0.661
rs1362349	45722	102573372	C/G	0.41	0.40	0.762
rs3755278	45967	102573617	A/G	0.07	0.07	0.984

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p-Value
rs3771180	47367	102575017	A/C	0.88	0.87	0.812
rs3771179	47642	102575292	A/C	0.07	0.07	0.868
rs985523	48126	102575776	C/T	0.16	0.13	0.270
rs1041973	49218	102576868	A/C			
rs3214363	49274	102576924	-/A			
rs873022	49433	102577083	G/T	0.57	0.56	0.868
rs3771177	49610	102577260	A/C	0.29	0.29	0.988
rs3732129	51282	102578932	A/G	0.51	0.52	0.795
rs1420101	51466	102579116	A/G	0.60	0.61	0.864
rs12905	53757	102581407	A/G	0.26	0.26	0.994
rs3771175	53960	102581610	A/T	0.82	0.80	0.444
rs3821204	54031	102581681	C/G	0.22	0.23	0.769
rs2160203	54574	102582224	C/T	0.23	0.25	0.732
rs1946131	55679	102583329	A/G	0.75	0.77	0.723
rs1054096	56100	102583750	C/T	0.66	0.65	0.807
rs2287038	56182	102583832	C/T	0.97	0.00	
rs1921622	59817	102587467	A/G	0.44	0.46	0.472
rs1861246	60533	102588183	A/G	0.23	0.24	0.824
rs1861245	60656	102588306	A/G	0.36	0.34	0.590
rs3755276	72209	102599859	A/G	0.48	0.51	0.423
rs2287037	72778	102600428	A/G	0.55	0.54	0.941
rs1420099	74293	102601943	C/G	0.58	0.58	0.904
rs3771174	77335	102604985	A/G			
rs1420098	78029	102605679	A/G	0.30	0.31	0.827
rs1362348	78374	102606024	C/G	0.04	-0.01	
rs1882348	78421	102606071	A/T	0.31	0.31	0.968
rs1558627	78434	102606084	C/T	0.60	0.59	0.839
rs2058622	79174	102606824	C/T	0.56	0.55	0.961
rs3836110	79397	102607047	-/G	0.70	0.69	0.643
rs3771172	79562	102607212	A/G	0.24	0.22	0.675
rs3771171	79700	102607350	A/G			
rs3771170	79730	102607380	A/T	0.26	0.27	0.700
rs2160202	79904	102607554	C/T	untyped	0.64	NA
rs2058623	79920	102607570	A/G	0.65	0.62	0.389
rs3771167	79938	102607588	C/T			
rs3771166	79972	102607622	C/T	0.53	0.53	0.820
rs1974675	80125	102607775	C/T	0.55	0.56	0.842
rs1465321	80368	102608018	A/G	0.29	0.30	0.781
rs2041740	83484	102611134	C/T	0.28	0.30	0.658
rs3771164	85536	102613186	A/T	0.73	0.74	0.905
rs2270298	85829	102613479	C/T	0.19	0.18	0.654
rs2270297	86425	102614075	A/G	0.57	0.53	0.249
rs2041739	88083	102615733	A/G	0.42	0.41	0.892
rs2080289	88770	102616420	C/T	0.61	0.60	0.840
rs3821203	90622	102618272	A/G	0.62	0.63	0.927
rs3771162	90924	102618574	A/T	0.26	0.25	0.621
rs3213733	91634	102619284	G/T	0.75	0.74	0.728
rs3213732	92029	102619679	C/T	0.39	0.45	0.176
rs1035130	95152	102622802	A/G	0.62	0.63	0.792
rs3752659	95348	102622998	C/T	0.79	0.80	0.826
rs3755274	96145	102623795	C/T	0.27	0.29	0.618
rs2241117	96793	102624443	A/G	0.70	0.67	0.480
rs2241116	97015	102624665	G/T	0.15	0.15	0.849
rs881890	97064	102624714	C/T			
rs3771161	97711	102625361	G/T	0.68	0.70	0.681
rs3771160	97855	102625505	A/C			
rs3771159	98708	102626358	A/G	0.37	0.37	0.970

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p-Value
rs1420104	not mapped	not mapped	C/T			
rs2041738	not mapped	not mapped	A/C			

[0233] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotype results. These values are indicated in bold. The allelotyping p-values were plotted in Figure 1 for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1 can be determined by consulting Tables 13. For example, the left-most X on the left graph is at position 102527857. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0234] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than 10^{-8} were truncated at that value.

[0235] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is place at the 3' end of each gene to show the direction of transcription.

Example 5

Gene expression profiling in IL-1 beta and PMA stimulated SW1353 cells

[0236] The human chondrosarcoma cell line, SW1353, (ATCC HTB-94) was grown in L-15 media containing 10% FCS. Culture conditions were at 37 degrees with 0% CO2 with media changes every 2-3

78

days. SW1353 cells were grown to 80-90% confluence in 10 cm dishes and then stimulated with either 10ng/ml IL-1 beta (human recombinant, Research Diagnostics) or with 200nm PMA (Sigma). IL-1 beta stimulation was for 3 and 24 hours and PMA stimulation was for 3 and 24 hours. Control cells were grown and extracted in parallel with treated cells. As shown in Figure 6, IL1RL1 was seen to be upregulated by IL1-beta and by phorbol esters in a human chondrocyte cell line model (SW1353 monolayer cell line).

[0237] The expression profiling in IL-1 beta and PMA stimulated SW1353 cells grown in 3-D alginate cultures W1353 cells were cultured as above and then resupended in 1.2% alginate beads at a density of 4 millions cells/ml according to the manufacturer (Cambrex). Cells were grown for 2 weeks and an alginate bead was removed from culture and tested for the presence of proteoglycans by Alcian Blue staining (Sigma). Positive staining indicated that the chondrocytes were expressing ECM proteins. Alginate cultures were then stimulated with IL-1 beta for 24 hours or with PMA for 3 hours. Control cells were grown and extracted in parallel with treated cells. As shown in Figure 7, IL1RL1 was seen to be upregulated by IL1-beta and by phorbol esters in a human chondrocyte cell line model (SW1353 3-D alginate cell line).

RNA extraction and cDNA synthesis

[0238] Cells from control chondrocytes and stimulated chondrocytes were isolated at the appropriate time period. mRNA was isolated from total cell lysates using poly dT beads according to the manufacturer (Dynal). Isolated mRNA was used to generate cDNA using SuperScript II reverse transcriptase according to the manufacturer (Invitrogen).

Expression profiling using semi-quantitative PCR

[0239] cDNA levels were normalized using the housekeeping gene, GAPDH. Specific primers corresponding to MMP8 and MMP13 were used in semi-quantitative PCR as positive indicators of induction of an osteoarthritic phenotype. All specific primers used, including MMP8, MMP13, BVES, CHDC1 and IL1RL1 (transmembrane form, soluble form, soluble isoform 1 and soluble isoform 2) for semi-quantitative PCR and are listed in Table 16.

TABLE 16: Primer Sequences for Expression Profiling

Gene	Forward primer	Reverse primer
GAPDH	ATCATCTCTGCCCCCTCTG	GAGGATTGCTGATGATCTTC
MMP8	CAATACTGGGCTCTGAGTGG	GGAAAGGCACCTGATATGC
MMP13	ATATCTGAACTGGGTCTTCC	GACAGCATCTACTTTATCACC
BVES	AACAGTATAGCCAGCTCC	ATCATCATCTTCTGCTCC
CHDC1	CCAAAGATCAGGACATGGATA	TGCTGTTTGTGGTAGGAGAG
IL1RL1 (TM)	CCACTCTGCTCTGGAGAGAC	GCCTGCTCTTTCGTATGTTG
IL1RL1 (Sol)	TCCGTCACTGACTCCAAGTT	TTGCTGCTGTGGAATACATG

Gene	Forward primer	Reverse primer
IL1RL1 (ST2_3)	AGGCTTTTCTCTGTTTCC	GTTGAATTCTTGGTTCACC
IL1RL1 (ST2_2)	TAATGTGATGACTGAGGACG	TGCAGAAACTCTGACACC

[0240] In a human chondrocyte cell line model, IL1RL1 was seen to be upregulated by IL1-beta and by phorbol esters. IL1RL1 has an unknown function, but it may possibly mediated inflammatory responses that can contribute to the development of OA. IL1RL1 is druggable by antibodies or by protein agents.

Example 6

In Vitro Production of Target Polypeptides

[0241] cDNA is cloned into a pIVEX 2.3-MCS vector (Roche Biochem) using a directional cloning method. A cDNA insert is prepared using PCR with forward and reverse primers having 5' restriction site tags (in frame) and 5-6 additional nucleotides in addition to 3' gene-specific portions, the latter of which is typically about twenty to about twenty-five base pairs in length. A Sal I restriction site is introduced by the forward primer and a Sma I restriction site is introduced by the reverse primer. The ends of PCR products are cut with the corresponding restriction enzymes (*i.e.*, Sal I and Sma I) and the products are gel-purified. The pIVEX 2.3-MCS vector is linearized using the same restriction enzymes, and the fragment with the correct sized fragment is isolated by gel-purification. Purified PCR product is ligated into the linearized pIVEX 2.3-MCS vector and *E. coli* cells transformed for plasmid amplification. The newly constructed expression vector is verified by restriction mapping and used for protein production.

[0242] *E. coli* lysate is reconstituted with 0.25 ml of Reconstitution Buffer, the Reaction Mix is reconstituted with 0.8 ml of Reconstitution Buffer; the Feeding Mix is reconstituted with 10.5 ml of Reconstitution Buffer; and the Energy Mix is reconstituted with 0.6 ml of Reconstitution Buffer. 0.5 ml of the Energy Mix was added to the Feeding Mix to obtain the Feeding Solution. 0.75 ml of Reaction Mix, 50 μl of Energy Mix, and 10 μg of the template DNA is added to the *E. coli* lysate.

[0243] Using the reaction device (Roche Biochem), 1 ml of the Reaction Solution is loaded into the reaction compartment. The reaction device is turned upside-down and 10 ml of the Feeding Solution is loaded into the feeding compartment. All lids are closed and the reaction device is loaded into the RTS500 instrument. The instrument is run at 30°C for 24 hours with a stir bar speed of 150 rpm. The pIVEX 2.3 MCS vector includes a nucleotide sequence that encodes six consecutive histidine amino acids on the C-terminal end of the target polypeptide for the purpose of protein purification. Target polypeptide is purified by contacting the contents of reaction device with resin modified with Ni^{2+} ions. Target polypeptide is eluted from the resin with a solution containing free Ni^{2+} ions.

Example 7

Cellular Production of Target Polypeptides

[0244] Nucleic acids are cloned into DNA plasmids having phage recombination cites and target polypeptides are expressed therefrom in a variety of host cells. Alpha phage genomic DNA contains short sequences known as attP sites, and *E. coli* genomic DNA contains unique, short sequences known as attB sites. These regions share homology, allowing for integration of phage DNA into *E. coli* via directional, site-specific recombination using the phage protein Int and the *E. coli* protein IHF. Integration produces two new att sites, L and R, which flank the inserted prophage DNA. Phage excision from *E. coli* genomic DNA can also be accomplished using these two proteins with the addition of a second phage protein, Xis. DNA vectors have been produced where the integration/excision process is modified to allow for the directional integration or excision of a target DNA fragment into a backbone vector in a rapid *in vitro* reaction (Gateway™ Technology (Invitrogen, Inc.)).

[0245] A first step is to transfer the nucleic acid insert into a shuttle vector that contains attL sites surrounding the negative selection gene, ccdB (*e.g.* pENTER vector, Invitrogen, Inc.). This transfer process is accomplished by digesting the nucleic acid from a DNA vector used for sequencing, and to ligate it into the multicloning site of the shuttle vector, which will place it between the two attL sites while removing the negative selection gene ccdB. A second method is to amplify the nucleic acid by the polymerase chain reaction (PCR) with primers containing attB sites. The amplified fragment then is integrated into the shuttle vector using Int and IHF. A third method is to utilize a topoisomerase-mediated process, in which the nucleic acid is amplified via PCR using gene-specific primers with the 5' upstream primer containing an additional CACC sequence (*e.g.*, TOPO® expression kit (Invitrogen, Inc.)). In conjunction with Topoisomerase I, the PCR amplified fragment can be cloned into the shuttle vector via the attL sites in the correct orientation.

[0246] Once the nucleic acid is transferred into the shuttle vector, it can be cloned into an expression vector having attR sites. Several vectors containing attR sites for expression of target polypeptide as a native polypeptide, N-fusion polypeptide, and C-fusion polypeptides are commercially available (*e.g.*, pDEST (Invitrogen, Inc.)), and any vector can be converted into an expression vector for receiving a nucleic acid from the shuttle vector by introducing an insert having an attR site flanked by an antibiotic resistant gene for selection using the standard methods described above. Transfer of the nucleic acid from the shuttle vector is accomplished by directional recombination using Int, IHF, and Xis (LR clonase). Then the desired sequence can be transferred to an expression vector by carrying out a one hour incubation at room temperature with Int, IHF, and Xis, a ten minute incubation at 37°C with proteinase K, transforming bacteria and allowing expression for one hour, and then plating on selective media. Generally, 90% cloning efficiency is achieved by this method. Examples of expression vectors

81

are pDEST 14 bacterial expression vector with att7 promoter, pDEST 15 bacterial expression vector with a T7 promoter and a N-terminal GST tag, pDEST 17 bacterial vector with a T7 promoter and a N-terminal polyhistidine affinity tag, and pDEST 12.2 mammalian expression vector with a CMV promoter and neo resistance gene. These expression vectors or others like them are transformed or transfected into cells for expression of the target polypeptide or polypeptide variants. These expression vectors are often transfected, for example, into murine-transformed a adipocyte cell line 3T3-L1, (ATCC), human embryonic kidney cell line 293, and rat cardiomyocyte cell line H9C2.

[0247] Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the claims which follow. All publications or patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[0248] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents. U.S. patents and other publications referenced herein are hereby incorporated by reference.

Nucleotide and Amino Acid Sequence Embodiments

[0249] Table A includes information pertaining to the incident polymorphic variant associated with osteoarthritis identified herein. Public information pertaining to the polymorphism and the genomic sequence that includes the polymorphism are indicated. The genomic sequences identified in Table A may be accessed at the http address www.ncbi.nih.gov/entrez/query.fcgi, for example, by using the publicly available SNP reference number (e.g., rs1041973). The chromosome position refers to the position of the SNP within NCBI's Genome Build 34, which may be accessed at the following http address: www.ncbi.nlm.nih.gov/mapview/map_search.cgi?chr=hum_chr.inf&query=. The "Contig Position" provided in Table A corresponds to a nucleotide position set forth in the contig sequence (see "Contig Accession No."), and designates the polymorphic site corresponding to the SNP reference number. The sequence containing the polymorphisms also may be referenced by the "Nucleotide Accession No." set forth in Table A. The "Sequence Identification" corresponds to cDNA sequence that encodes associated target polypeptides (e.g., IL1RL1) of the invention. The position of the SNP within the cDNA sequence is provided in the "Sequence Position" column of Table A. If the SNP falls within an exon, the corresponding amino acid position (and amino acid change, if applicable) is provided as

well. Also, the allelic variation at the polymorphic site and the allelic variant identified as associated with osteoarthritis is specified in Table A. All nucleotide and polypeptide sequences referenced and accessed by the parameters set forth in Table A are incorporated herein by reference.

TABLE A

RS_ID	Chromo -some	Chrom Position	Contig Accession No. [1]	Contig Position	Nucleotide Accession No. [2]	Sequence Position	Amino Acid Position	Locus	Locus ID	A [3]	Allelic Variability	OA Assoc. Allele
1041973	2	102576868	Hs2_22327_34:13	5021492	NM_003856	coding-nonsynon	E78A	IL1RL1	9173	R	[A/C]	C

[1] Contig Accession Number which can be found in the NCBI Database:
http address: www.ncbi.nih.gov/entrez/query.fcgi

[2] Sequence Identification or Nucleotide Accession Number which can be found in the NCBI Database:
http address: www.ncbi.nih.gov/entrez/query.fcgi

[3] "A" column is the sequence orientation ("F" is forward, "R" is reverse).

[0250] The following is a genomic nucleotide sequence for an *IL1RL1* region. The following nucleotide representations are used throughout: "A" or "a" is adenosine, adenine, or adenylic acid; "C" or "c" is cytidine, cytosine, or cytidylic acid; "G" or "g" is guanosine, guanine, or guanylic acid; "T" or "t" is thymidine, thymine, or thymidylic acid; and "I" or "i" is inosine, hypoxanthine, or inosinic acid. Exons are indicated in italicized lower case type, introns are depicted in normal text lower case type, and polymorphic sites are depicted in bold upper case type. SNPs are designated by the following convention: "R" represents A or G, "M" represents A or C; "W" represents A or T; "Y" represents C or T; "S" represents C or G; "K" represents G or T; "V" represents A, C or G; "H" represents A, C, or T; "D" represents A, G, or T; "B" represents C, G, or T; and "N" represents A, G, C, or T.

IL1RL1 Genomic Sequence (SEQ ID NO: 1)

```
>2:102527651-102626600

1      tgcaggattt gtatgtaagt atccaacaga caacacctgg ggtgatcctg tgaatgtaga
61     ggaagacgtg atcttaaatg gaccagggaa gtaagacctt aggatgtacc agccagctga
121    gtggcaggta atattgattc tcatgtttca ccccatgttc ggtcacttgt gagcagattc
181    ttcagttcaa ttctctggct agttaaYttg gtctgggagt cacaccagaa aatgtttctg
241    aaaataggta cctattgtga aaagtataat cctctttgct attgtgagta gtgctgcaat
301    aaacatacat gtgcatgtgt ctttatagca gcatgattta tattcccttg ggtatatacc
361    cagtaatggg atggctgggt caaatgatat ttctggttct agatccctga ggggagcggg
421    gagggatagc attaggagat atacctaatg taaatgacga gttactgggt gcagcacacc
481    aacgtggcac atgtatacat atgtaacaaa cctccacgtt atgcacatgt accctagaac
541    ttaaagtata ataaaaaata tatataaagt ataatcatca aattgttaaa atcatgattc
601    tcatgaagta gattgtgtct gttgaagttt aattaatgaa ggaaacaaca ggatgtcagg
661    actgaatcca taaggatctg agaaactggt atttccataa gatgttaaga taatattgat
721    aactgggata caaaactgct taatgtccta caatgtaata aaacaatttc tgttgtgttt
781    tgcatgagat agaaacaaaa atacctccca ttggtcagtt ttctacaaga taatctttaa
```

```
 841    ctttacaggg ttgtaaaata gctatgcctg aattgttagt taaaacccca tttttttttt
 901    tagagaaaac aatgatcctt agactcactt actagaaata tgttctagtt tttttataaa
 961    aataaaataa aataaaatatt tttattataa aataataaat atttttataa taaaaataaa
1021    aataaaataa aaaatattta ttttaaatat tttttttcaag catcagtata ccttattctt
1081    gttcaacttg gggaatggcc ttattcaaat tgggaatcct cagctcatct ttgtctttgg
1141    cattccagag aatgaatgtg tttgggaaat tgtgtcagct gagcagtgta tacagaaagg
1201    gacaggagac cactgaaatg gttcaggtta tagttagata gttcctgcca ctggttagat
1261    caggagagtt agatggttat agttagatag ttctttatgc tggcagtggg aaacaagaga
1321    agaggacata tgagggagat gctcatgtta ttgggaacaa tatttggctc ctggtagaat
1381    gttaggggcg agggagaaga ttctaatgtg ttgagttcgg gagccgctgc gccattaata
1441    aaaccagaaa ccttctacag tgctccaggt gggagccatt agcatcaaag gcgattctaa
1501    aatcagtcct gaaagctatt ctatcccttg ttctgcatgc agaacagctc ctgagggcct
1561    gttgagggcc cagggaaact ttggcccagg gccaaggaag gactggtgtt cccttttggt
1621    ttagctctgc tcacagacaa ctgcccagaa ggtccaggtt ccaagtcact gtgcaagctg
1681    acacgtgtga gacataggaa gctgttggga ctacagaacc cagggactcc aaaccctgtt
1741    taagtaccac acctgagggg aagaagaccc tgcttcccaa ttagtcagct gggtagaagg
1801    ctgaggattt ttgaaaagag gtagttggtt tagcctcagc tttctttaga ataatattgt
1861    ggtttgaaag tgtgcaacct cctagtgtca tcatttcctt ttgtccctga cagaactgtc
1921    ttcacgctgg aaataaaagt atgggacagg cttgaagcag agttgtgaag tagaagtttg
1981    ctaggaatcc attttttatc gaaagcccca ggcaactctg attaatttgg tctgggagtc
2041    atacatggag aaatactgtt ctaggagctc attcagctaa cacctctctc acctagaggt
2101    catgcgggga aagaaatgtg tattgcatta taattcatag tagactgtga ataactaata
2161    ggaaaagatt cctttcagct tctcctgtct agttaagttc cagcagggac acacctttag
2221    cttaaaggaa tgcattttag tggtgatggt ggtggtggtg gtggtgatgg tggtggtggt
2281    ggtggtggtg gtgtgtcttc ctaggtcctc tcaaatcaca tcacatggtg gagggataaa
2341    aagaggcagg taaaaaatag catttgttct acaattgatg aatgactaaa tgtcctcttg
2401    aacataaaga atttatattg gtaaaagctt ctctgctatg tgcattagta aaaacaacaa
2461    catcgtgccg ggcgcggtgg ctcacgcctg taattccagc gctttgggag gccaaggcgg
2521    gtggatcaca aggtcaggag atcgagacca tcctggctaa cacggtgaaa tcccgtctct
2581    actaaaaata caaaaaatta gccagcgtgc tggcaggtgc ctgtagtccc agctactcgg
2641    gaggctgagg caggagaatg gtgtaagtga acccgggagg cggagcctgc agtgagccga
2701    gatcaggcca ctgcaatcca gcctggggga cagagcaaga ctctgtctca aaaaaaaaaa
2761    tcgtttgctg atttagaact cttcttaaag agcctccata tgtttaaaga gccatgcctt
2821    atgactgtat gactataatt ttataaagcg ctttccctac agtatcttac taaatccttt
2881    cagcaatcct ggaagtacca tcattatctt tattttgcat attcaagttc tgaaagcctg
2941    gtttgctcca gtaagtagta gagctaggac caaaacccaa gttttggttg ggttttcagc
3001    cctgaatctt tcagccctga accttttcat agctgcctct gacttcacag ctcactgcat
3061    aaagccttct tcagacccat cctctgtcag catcgtctag tctagcctgt caggtatttt
3121    ggcctccaca ggagatccta ttctcaatgt ttgtaaatat ttagctaaat ctgtagcaga
3181    catctggttt cataatgaga aagagcattt ttgttttcat tagggaaaaa gagcatgact
3241    gggcgcagta actgcagttt tcatcgtgga gaaaaagaga gtgagacttg aacttgaaaa
3301    agtgaagcca agcagtgcag ttggcaggtg gagggtgggg aggagggaga ataaagacaa
3361    gtcaggagga gagagaggaa aatgatgcac agagaaagag ggaagaggca aagagaaagg
3421    aagggagcct ggacagcaga ggggtctcct tcaggaaaca gagttgaagt gaaatagtgt
3481    aatccccagg tggcaggtac aacatttctc agcacaatga ctctgaacaa cttaattctt
3541    tcttagggga agtagttaca cggaaagtgt aatgtgtgct gcacatttgt agttgtgata
3601    ccatcagaaa gaggtacttg ctgtggaaag aacttctcaa cccctccatg taatgaccct
3661    tggagaatgc attatcaaag cagagtcctg cagcctgaaa gtaaacagca tctacttaga
3721    aatacagtga catgtaaact ttccccaaat caacagtgaa accagtagca ataactttga
3781    tgaaacaaag gggcattgcc gagctgtcag gacacaacag agcacaacag aacagcccca
3841    ggtaaatatt gataaattag gttgatctgt gtattcttag tcatgggaag ctggggcaga
3901    taacgaaacc ttgtatgtgt cctttatagc tctgtctgat gagtttgggc gacggagtct
3961    tgctctgtcg cccaggctgg agtgcaatgg cgtgatctca gctcactgca acctctgcct
4021    ccctgattca agtgattctc ctgcctcagc ctcctgagta gctgggatta caggcacgtg
4081    ccgccacgcc aggctaattt ttgtattttt agtagagaca gggtttcacc atgttgatca
4141    ggctggtctc gaactcctaa actcatgatc cgcccacctc agcctcctga gtgctggga
4201    ttacaagcct gagccaccgt gcccagcctt gttttttttt aagtgatttt tgagggcagt
4261    tttatgggca gaaaagactc ttcgtgaatt tattttgatt ttaaaattac tattcaattt
4321    ttcaaagtta aaaatgtaaa aaatctagtt tattataaat tgaagaccag cttaaaaatc
4381    atataatacc atttatatgt ctagcacatt ttaacaaact tcaacactag gttactccac
4441    ttacaaattt ggcaaactaa atattttaag gtattttaaa ccgcatttat acatttactg
4501    tatttgattt ttcttcttaa gaactttggc taaaaccttc ccaagtaatg aattaattct
4561    tataaattta agttctctta gaaattttga cgctacatat aagcttaaat attggatata
4621    aacttagtaa ggatttcact taaaatcgta aatctaaaca atttccaatt ataccttttt
```

```
4681    aaaaaagagt tacatgactt attctaatgg gtcaatattt taggcatcac aaatacataa
4741    ttgcaaatat atagattcct taaattatac catggtcttg attcttatta catagatgtg
4801    tacatacctt tctacccagg tctaaatacc actgaattca aaatgagtct tgcattctaa
4861    gtacttcgaa attcttaatg ggtattaaaa agtgacccag ttgcttatgg ataaacctca
4921    taatttgcct tgatacttgg tcacaaatta aaggtcattc tactgacaaa ttacttgtat
4981    acaatctaca aaacaaatag catgcaattc tcattacagc acacgtgttc agactttaac
5041    aaagctgttt gaagtctaca tcatctcctg agttagagaa gccccaattg ccacatacag
5101    atacttagta ctgtcataca gagtcacata tctactggat tctacctgga tcttaggatc
5161    aagatgcttg tactcagaca aacaaattag gaaagtagcc tcctttgcag agtggtctag
5221    gggtgttacg agccaaaaaa tcattcacgg ttttatggtt aagaagtgag atgatctgta
5281    atgaaacctc acagagattc tgttaaataa caagtgccca aaagctactt ggaaaaggag
5341    tagactgaag agaaaatgag agaatcaaaa atagttttaa tttcctgtta gaatcatctc
5401    ctttacttta ctcttgcatg tgtaactgtg gaggagattg taaacattcc cttagctaac
5461    tgtacttgtt actcttgatc aacatccctt agtaactgtg tatatgtgat gcttcatgca
5521    ttagtcagag ttgagaacaa aaacagagta gctcaaacaa aacataggtt tatctgtccc
5581    tcatgtaaca acacatgggt agccaggtgc tctagggatg gtagatggat ctgcttcatg
5641    aagtcactca gaatccccaa ttcctcctac cttgctgctc tgccctgtgt ggtcctcatc
5701    tgtctcaccg ctgaagccac cttgcagttc acctagaaga gaatatgtac tgtttctgcc
5761    aaatctgtgc tatagatact agatgctttt aatggagagg ttagccaaga aacatgacat
5821    ttctcttata ggtaataaaa tacttctctc taaagaaata cttagtttaa aattttattc
5881    tgctatagtt ttagcagtag gataaaagct tagtttgcag aagcggaaat tatcctatat
5941    ttaggcaaag tgagctatct gtgtgttatg attagagaga ataccaagaa gtgatttcct
6001    tccttaaaat aacttctcYa aatttccacc attaactacc cattcttcaa ctctctcaat
6061    ttttcttatt cctcccatgg agggagtgca tgtatttaat tactaacttt actagtatta
6121    tataatctat gctccaaaat ggttagtcta caattagtag agggataagt ctttagggtt
6181    ggaaagaaaa aacagttatt aaagatctac tattgctgga tgtcatacta actcttctta
6241    cttattatgt tatttatacc acctaacagt ttgataatat tacccatccc cccgctttttt
6301    tttttaagag ataaggcata tatgacttat atagagtttta ggggatttttc ctcaagatta
6361    gacagttaga agtctcttgg gtgggagagc agttctaggg gaaaaggaaa gatMagaaag
6421    acaagtcaac tactttgagt aaagacctct ggctgagcag gtcattgtgt attcccagaa
6481    tttgtcacca actagttgtg tggttttgga catacccctt cacatcacct caatgtcacc
6541    tttggtaaag tgaaagaatt ggttgtagat attctttaaa ttatcttta ttcaggcttt
6601    ctacagccca attctctcaa ttgttgactt ttacttagaa ttttgggcaa acactagctg
6661    atacaatcca tttaccaaga tgtttatctg atgacatttt catcaagttt ctaaagatgc
6721    catattgtct ttctggcagg ctttttggaaa gtgttcaaga tctcctgttt tattttcaca
6781    gttggcccct gagcctcaag tgtaagcaga gctgggctga gaggcttagg tcctcctggc
6841    tgatgtcact aggctgaggg gccttgggga agaaaacaaa tttctcccag attcaaccct
6901    ccaaccaact gcaggtcctg gaacagtgag taactccttt actcagaagc tgaacattaa
6961    aagtcttttc tgattcttac caacttggca cataagaaat acagtcattt cctcaaatac
7021    attttttatta tgagggtact gaacattgta tttattggtc atttgcattt tttgtaaatt
7081    gtagattttt aaaaacctct ttgtctgggt gcggtggctc acacctgtaa tcccagcact
7141    ttgggaggcc gaggaaggtg gatcatctga ggtcaggagt ttgagaccag cctgaccaac
7201    atggtgaaac cccatctcta ctaaaaatac aaaaattagc tgggcttggt ggcgtgcgcc
7261    tgtaatccca gttactccgg aggctgaggc aggagaatag cttgaaacca agaggtggag
7321    gttgcagtga gccaagatcg Ygccactgca ctagcctggg caacaagagg gaaactccat
7381    ctcaaaaaac aaaaacaaac aaacaaacaa aaacattttt gtattaacta gtaaagtctc
7441    ctgatacatt aagaatatta aagacaggca tatagtgagg ttttttttgtt tagtttccat
7501    atgccttta atttgtttaa ggatatagtc atttgtttag actataggta aggaatagtt
7561    atatctttg ttccagaaag tttcagattt tgagggaata ttactattcc taaggtttta
7621    ttgacatata aattttaaaa atattatgta atgagagcta tatatatata ttctttttttt
7681    gtgcatttta tcatgggtac taggttaagt gtctgtttct atgaacatac aacttactta
7741    acatatgact tggaaacaga acatgcatag taaatgtttc tcaagtaaat aagtaaaagt
7801    atatgagaaa caaatgtgct gtttacatgt tgtcatagga atcatcttag cactaaatgg
7861    ataattagga tccaggttca tccttgggct tggcacaaat gagcttggac aatctgaaat
7921    atggatggtt gataataggt aattaggttt aaattatacc aggaaactga agaacattca
7981    gaaaatggtt acagataatt acagaaggtt gcattacaga ggaaaaattt ctcacagtga
8041    gagacaattg gcctcacatg cctttactga ggcatacatt gtgtcataat ttcctttttat
8101    tgcaccttca acattttgaa gggaattaat tgtctgaatt aggctagatg tcattttttt
8161    cttggaatta gaagaaagat tgagtaaccc cttatgtttta tctggaacat gtctttctgt
8221    gcatttaaaa atctcagagc ccttcagaac cttgcccgtg acccgtcatt gggaatatct
8281    taaaaatgtg tgtgtgtggg tgggtgtgtg tgtgtgaata taggaatgag ggaacactgt
8341    cgtagacaga tatatctgaa gcttatcttt tggtttaggg catttggtct tgcatcacct
8401    tttaagtctt atgctatcct tgactttttga acagaacaaa cttctagatt tttgtatgtt
8461    ctaagttctg cctttgcagt attagaggtc tattttccca gactagggac tggagtgttt
```

85

```
8521     aatcatcact gtcatgtaga cctaatatag aaatacaaat gcaataagaa aaatgttgag
8581     aaattaaagt cagaaacata tttgttgtga tacatgttca aaaataatat atttcattgc
8641     aggaaatagc aaatccagag aaataaaaag aaagaaaaaa aataatctgc atttctacgc
8701     atgcagctaa ccattgctga gattttggag tccttgtttc ctaggcacgt tttcctatgt
8761     ctgtgtgcac acatgtatgg gatgttgcat atgtgtatgt gtgcatgtgt ttacacatgt
8821     atttgcaggt atgtgttgat tattatacat tgagtcataa tgtatattac atttaaaatt
8881     tcctgttttt acggctttat taaggtataa ttgagatatg aaaaattgca tacacttaaa
8941     ttgtacagtc taagttttgt catatgtata cacaccagtg aaacccttac cacaatcaag
9001     ataatgaaca tacccatccc ccgtaaagct tctacacccc tttcgaagtc catccttcaa
9061     cttctcctac cgaatccaca ggcatccact gacctgattt ctgttactat aagtttgttc
9121     acattttcta gaaattgata taaatggaat aatacagcat gtgcttgttt ctctctggct
9181     tcttctattc tgcgtaatta ttgtgaaagt catccatatg attgcattta tcagtacatt
9241     catactattg catttatcaa tatgtcattc ttgttgattg tttatcatat tcccctatta
9301     tcttcttaac atctatagaa acttagtggt gtcaactctc tttttcttga tattagtaat
9361     ttgtgtcttc tttctttttt cctcagtcag tgtgactaga gagttatcaa ctttattagt
9421     cttctcaaca aaccagcttt tggcatcatt gatttttctt tattgttttt ctgttccatc
9481     tcactgtttt ctgcattgat ttgtattatt tcattttctt ctgttttctg tgctttcttt
9541     gctcttcttt ttctagtttt ttaaggtggt aaccaatatt gatttatgac ttttattatt
9601     ttctaatgta tgcctttagt gttgtaaatt tccctcttta gtggcatatc ccaaattttg
9661     atatgtagtg ttttttgtttc catttagttc acaatacttt ctaactgccc ttttgatttt
9721     ttgtttgcac catgggttat ttaaaagtgt attatttagt ttttgcacat ttacaaagtt
9781     tctatatatt tcttttattg attttgtgat ttaattccac tgttgtcaaa taacacactt
9841     tgtatgcctt gatccatttta acttattgag acttggtttt gggactggaa tatgctctct
9901     catggtaaat gatccttgta tgcttaagaa gagtgcttat tccagtgctg tcaggtggag
9961     agttctataa ataacaatta ggtctacttg cttaacaatg gttttcaaat tttctatatc
10021    tttacagatt tttctgtttc ttctatcaat tactgaaaga agaatattga catctctgac
10081    tgtaactcaa gatttgtcta tttctccttt cattttatc agtttttccc tcatgtattt
10141    tgaagctcaa ttattagaag cataaacagt attgttatat tcttcagata aattgacccc
10201    tttatcatta tggaataact cccttttattc ctggttatat ataatatatg gaataactct
10261    ctttattcct ggttatacat aatatatgga ataactccct ttattcctgg ttatatataa
10321    tatatagaat aactctcttt attcctggtt atgtataatt ttagagcaat atataatatg
10381    catatttttt gtctaatctt attatagcca ctccagccac gttttgattg gcattaatat
10441    ggaatatctt tttttctttct tttgctttga aacttgttgt gtctttgtat ttaaagttgg
10501    attcttgttg gcaaaatgtt gttgggtttt cctttttaaa tccaatctga gaatctctag
10561    cttttagttg gggtgtttgg agatcattta tgttaatgtg attgatagtt tgggcttaaa
10621    attcttctct tgctatttgt tttctatttg tcccatctgt tctttgttcc tcatttatcc
10681    ttttgtgtc ttattttgga ttacttgttt tttttataa ttccaaattt tctccattgt
10741    aagatcgttc tctatagctc tttgttgtga tattgtagtg gtttctctag ggttatctgg
10801    agttagtata catctttaaa taacacagtc caccttcaag tgatattgta ccaattcact
10861    gataatttaa gagctttaca acagtgaact tccatttctc ttctctggtc tttgtgctgt
10921    tgctgtcata catgttgctt ctacacccct attataaaac ccacattata ttgtcattac
10981    tttYgcttcg aacaatcaat tatcttttaa ataatttaaa caataagaaa taaaagtctt
11041    tttatttacc ccatagatgc catttctggt gctctgcatc ccttcctgtg gatcagatct
11101    ctgtctggta tcagttcttt ctacttggag gacgttcctt tcacatgtct tgtagtactt
11161    gtctgttact gacaaattgc ttcagctgtt gtatgcctga aaaagttttt atttcgtctt
11221    cacttttgaa caatattttt attaagtaaa aaattctaag gtgacttttt ctttcagtca
11281    tttaaagatt ttattctgct tgcattgttt ctgatgagaa atctgatgtc accctctccc
11341    tagttcctct atagtaacat ggttttttct tttctcactt ttaagagttt ttctttatca
11401    ctatttttag tgtttgagtt tctgggtact tcagggtttt atacttttca tcaaatttgc
11461    aatattttgg cgttatttct tcacataact ttttttctgt ccttttgttc actcctttcc
11521    tcctggaact tcaattatac ttacattaga tgactcgaaa tttccccaca attcacagat
11581    gctgtatttt tttcagatct gttttctttc tgtattttat tttgcgtagt ttctattcca
11641    tacaatatat ttttcatctg aaatgtagct tttatctgta gatgttttat ttaggtcttt
11701    ttaaaatctg ccatatctct gcttaacaca atatttcata taggtttttt tttcacatat
11761    ggaatttgga tataatagct gttttcatgt ccttttctac ttattctatt tatctatgtt
11821    aattctttgc tagttttgat tgattgattc ttctctttat taagggtagt attttctgct
11881    ccttttttatg cttgaaaatt ttttattgta tgccagacaa tgttaatttt aacatattga
11941    gtgctgaata tttttctatt tctataaata tttgagċtgt gctctgaatt actgttatgt
12001    tgccctgaaa caagtttgac atttttgggt cttccttta agtttttttt atgtgtaacc
12061    agagtgacat ttagcctgtg gttaatttat actactattg agacaaaact catgagttat
12121    gtcattttcc attttgaata gtgggcacag gtactatttc tgaccctgta agagttgtgg
12181    taattgttcc ctctgatcgt tttgggtggt tctttctttg accacagatc acttcctcag
12241    aagcatgaac tcgtcagtac tgagctgagt aactgaagag ggcgctgcag atatctctgt
12301    agtgatctcc tcagtactct gtgaactcaa gtcaccttag cctttccaga Stcccagctc
```

```
12361   tgattcctca  tctcggagag  accactggac  ttcatctctc  ttcctcttct  ctttgccaga
12421   gttcagaaac  tctccaggca  gtaaccaggg  caattgccag  cctctttttt  tgcttcccat
12481   ctctcaaggt  tcaccgtcgt  ccactgcctg  atgcttatat  cttgaaagct  gttgattcat
12541   acgttttgtt  cagtgtttag  ttatttaagt  caggagggta  aatcttattc  ctattactct
12601   atcttggctg  gaagtagaaa  tcacaaaact  tgcctctaaa  caaagtatat  cacaaacatt
12661   ttctatttat  ttaaatatta  ttagtctaca  tcattttgga  tgactacaaa  gtattttatt
12721   atttaaaaat  accatatatt  aattaatact  ttatttttta  tatttcgttt  tttcctgcta
12781   ttgttacaat  gttgtgaatg  agcagataaa  tatttttgaa  tatccatgat  tatctcaact
12841   gtattagtag  aagagaaatt  gctgagttaa  ggggtatttt  gatgcagatt  tccaaattat
12901   cgtctaaaaa  ggttgcaccg  atttatatac  ttattatctc  tccaaatatc  tgggtattat
12961   tattataatt  tttttaaacc  tccaagcttc  ataagtaaaa  atggcacatt  tattggtttt
13021   aattagcatt  tttatttctt  ttgatgtata  atattttaat  tagtttattg  aaaatctctg
13081   tccttccttt  ttgaatttag  tttatatctt  ttgcatattt  taaagtttag  attgttgtct
13141   tttctgttat  taatttataa  aatatttgtg  ttgagtgtgt  gagtgcatgt  gcatgcatgt
13201   gtatgcgtat  gtgtgtgttt  gtaacatttt  catcagtttt  ttgcttgccc  tctatttatg
13261   gtattttcag  atgtgattag  atttatttgg  tctttgacat  tatgatttct  gtctctatca
13321   ttgagtttag  aaggYcaata  ttattcctac  ctcattacca  aagaaatact  taattctatc
13381   tgcttacaga  tttaactttt  ttcttttttcc  ttttaaatct  ctaatacatt  tggaatttaa
13441   ttttcttact  gattaaactt  tacttgtata  tgtttatatt  ttataatgat  aatattaaac
13501   catttgtaat  gcttaaatta  aatattaaat  ggtcataaca  ccatttattt  aaagcctatt
13561   ctttctttca  ggtttgatat  actgactttt  tcatcatgtt  gataatcttg  tgatatattc
13621   ctgttaatgc  tccattggtt  tctacatttt  ttatcctgtt  cttataagtt  gtctgtctcc
13681   tcctgtggct  gttctatttt  attttaatta  ctgaaacttt  attggacaat  agtaagctca
13741   gaagagaata  tctgtaggac  aacagtaggc  tcaaggagaa  atcatctgtt  ctcctaatct
13801   tttactattt  ttcgtatttt  cctgactgat  cttcagtatt  tctggaaaat  tgacaagcaa
13861   tgcagatttt  ggataatagt  gcaggcttca  ctttcattgt  gcctgggttc  aaatcccagc
13921   ttcttagctg  actggcaatg  tgacccagca  aatttatttta  ctttggcatg  gctcagtttt
13981   cttatctgta  aagtaagggt  aatagtagtc  cctatattat  aagtttatta  tgatgattgt
14041   gtgggtggaa  cccgtgtgtg  gcacacaccg  gcacttagtg  agaagttagc  tttgccattt
14101   tcctttattc  ttcactttaa  ttttagaatt  attttgtcaa  cttcaaaaaa  gttcattatg
14161   aatcttatta  ctgagtcttc  tctagaagca  tgacatgact  cttcattcac  ttgttttctt
14221   ttgcgtctct  cgggaagttt  tgcattttct  gttgtgtata  tataatattt  gtacattata
14281   tctaactatc  tatctatcta  tctatctatc  tatctatcta  tagatatata  tatatacttg
14341   cacaattata  tccacataca  ctctaagaaa  taaggccagc  agaaaatatg  tgggtatgtt
14401   agactataat  tatctctgag  tattaggtta  tgagaatttt  ttcatttgta  gtattctata
14461   cattttaaaa  cttattaact  gagcattact  gatataaatc  aataattttt  aagtgacttt
14521   tgtgccctac  tacgctggtg  ctgtgtgctt  cagactgaag  ggcagtgcct  gttagatctt
14581   atgtgctaca  tgttacacaa  gtactttggg  aaataagctg  ttgttttaaa  aaatgtgtca
14641   tcaatttcta  gagaactgga  cataactttg  aagcaaaaca  attttaagtg  gatgactaaa
14701   atccttcctg  aaaacgcatg  cctgtggtcc  ctgtgaaatc  acctttttctc  cccctttcct
14761   aatctccttt  ccctcctaaa  tctaaatttt  cttgtaatgc  ttcaaaatat  ctggacttga
14821   tctcagaaga  tattctgtca  tgcatattga  caggactggg  tactttcagc  aaaacaaagg
14881   acatctttaa  atttttacca  gccccagata  gaagtgtcta  gagttctaga  ttttaggaga
14941   aatacagcca  tgaaggaaaa  tgttgacatt  tatttcaacg  tttcaacagg  aggaaagaat
15001   gagaacacaa  tgtacagatc  agcctttcct  ctcatgttct  tacaatggga  gaaggagaga
15061   gggtatttga  gacaccaaag  catgaaagcc  atgacacagg  gaaaatactg  agaacataac
15121   ctacagatca  gattttcctt  ttatgctctt  accatgtcat  aaggcaaagt  gcttgtttga
15181   gatttcaggg  cacaagagcc  actacagcag  gaaggcgtcg  catcatcgga  acatttgtga
15241   tcaaagagaa  gactgtctgc  acaaagcacg  gactgcagcc  aagggtcct  tcacccactg
15301   ttctcacaaa  tgcttgcctt  tgctatcaca  actgctattg  ccagctggac  gctgagtgtc
15361   ccctcagggt  tcatagagat  ctgtctccct  gaggtctctt  aggtgctgct  gctgaccagt
15421   agtaacttgg  tgaagaccct  cgattgtgag  gacattgtgg  tgaaatgttg  ctttcttgac
15481   aaaacaaatg  tgtttccttc  tgatccatga  acaaagatct  gttgaataaa  tattatgagc
15541   gtgtagaagg  gagagacttt  atctgatcag  actagggaaa  gggtgctttt  aatatctgtg
15601   gaacaatctg  taagatcgga  cgtttcctct  gataacagaa  actccagagt  tacaagagaa
15661   atggcttcct  gacatcacag  acactaaaag  gccttctgag  taggagtagc  agttgaattt
15721   ttgactacta  tgtaagggaa  atacagtacc  aaaaaaaaaa  aaaacccatg  gagatcaagt
15781   catttcttac  tagcaagtac  ccctgacttc  tacgtatttg  ttgttagatc  cacaaagaat
15841   atagtataac  ttcattttct  atcaactatt  agggatctat  agtgccataa  caaagtactc
15901   taaaatttat  cagcttaaac  agtaaacctt  tattatctca  catggttcct  gagggtcagg
15961   aatctgacag  tgacctagct  tggtgattct  gcctcggggg  tgctcatgaa  gttgtggtca
16021   agctgttgac  tacactttt  catctcggaa  gacttgactg  caactgaggg  acttgctgct
16081   agggtcactc  atgtggctgt  tggcaggagg  cttcaggacc  tgtctgtgtg  gagcactctc
16141   tggggctgcc  ctcgagatgg  cttctctcag  aggggggtgat  tcaagagaca  gagtcccaga
```

87

```
16201   caagtacagg atgccttaca taacctaatc ttggaaatgg cacgtgatca cttattccat
16261   attatattga tgacatagac ctgacctggt acagtgaggg aaaagactac acaaggatat
16321   aaacaccagg atgcagagat caggggggcca tcttggaggc tgcctgccac agctattaaa
16381   ggtaatttgt tgacaactca caatttctag ctacccagtt gctcattctt ttgcttaaac
16441   tattcaaggc tggtggggga catgtgaatc ctgttaaaat gtggaaagat ctaaatgtgt
16501   tttatagaaa aaattgagta aacttttttt gcttaataag caaaagaatg atgaggaagg
16561   aaggggcaat tgtgtgtgtg tgtNagagag agagagagag agagaatata aaaatctaat
16621   tttttaccaa aattatttt tttcttccgg gtgaacatgt ggctttaaca taacttctat
16681   tccacaaact caaaatcaag cacctctgtt aactcttaac taaacataga cttttttStga
16741   actttggaat aggagtaagg gggaagagtg aaacacacta gctttcatgc cttgtttaat
16801   ttctaatttc ttaaagagtg gacaagaaag agaatgaatt cagcaaagtc tttgatccct
16861   acagttcata ttcagacaaa taggagccaa ttattaaggc taactgcctc agtaaaataa
16921   aaaacagcca atttaacttc tgaacacagg aaggtttttt acccattatg gaatcatgaa
16981   agatctgaat gcacatgata tggtgtaacc ataaaataag gttgagtttt tgcttcctgt
17041   taaatctcat gagaacaagc attctgagtg aggggtgcgg taaagacaca cagggctagt
17101   cttaaaagaa ggtaaagttt ctggctatgt attaggcttg attagaattt tctggttaaa
17161   acgaattctc cttcagtttt gtttctttca tagatgctgg aattcctata aaagccattc
17221   agacatttat atttatgact tttctacttt ttaaaaaact atactcctcg agagaataaa
17281   ggagtaggtt ttactagaca ccatatttaa tgggttaaag caaccaagta agaagaggga
17341   atgttctaga agtgagggaa aggggattca tgcagacatg aaagtgaagc aaaggtaaat
17401   gtgacagaga aaaatgaatt tgccatctct ttctttcttt ctctctttct tccttccttc
17461   cttccttcct tcctccctcc ctccctccct ccctccccct ctctctctct ctctctcttt
17521   cttttctttct ttctttctttt ctttctttct ttctttcttt ctttctttct ttctttcttt
17581   ctctctttct ttctctcttc tttccttctt tctttttttt agccacagtt tcgctgttgt
17641   cgcccaggct ggagtgcaat ggcaggatct tggctcactg cgacctctgc ctcccaggtt
17701   caagtgattc ttctgcctca gtctcctgag tagctgggat aacaggcatc tgccaccacg
17761   cccggctcat ttttgtattt ttagtaaaga cgggatttca ccatcttggt caggctggtc
17821   ttgaactcct gacctcaggt tatccaccca cctcggcctc ccaaagtgct ggcattgcag
17881   gcatgagcca ccatgcctgg actatctgcc ttctttctaa aagagataaa gctttctagc
17941   tgcctagtac cttggtgaat ttgggctgag gccattcagc gacaagcact ggaaagaatt
18001   gaagtgctcc caagtacctc tgcaagggaa gaggggtgat gagcagggga gaggagagga
18061   ggaagaggac ctggtggatg gtaggaggcc atcaggtctt ccatgtgatg tgcgaatctt
18121   tagctaggta caccgtattt tagaaatgtt tggtgtaaat ggatttcttt tgcaaactgg
18181   agaggggact gagtttttggg ctgatgctga gtggaacaga gatgggactc ctggagcttg
18241   agacacagga tcttggagaa aatgagcccc agaaatcagc agatgtcatg aggggactct
18301   gggttatcat aagctgtgga attggtatcg gatctgaact tgtttcttca aagtgcaagg
18361   aagaccagct aatctcttgg ttttgccaat ttagggcaat tttcaacatt acgaaaattc
18421   atttagtttg ggtgcagagc tcagatgtcc tctccatgtg gacacaaccc tacatgtgca
18481   cacacaaaca agcacttgag cacacatgtg tgtgcacaaa ccccactttt ctgcttttcc
18541   aggagcaccg tatgttattt cagacaaata agtttggcca actttgatct ttccattcag
18601   gtcactgatt gctgaattgc cacagatgtt gttctgccct ccttaggttc tagggcattt
18661   tcattaagga gagttacctc caagaggggt catggcaggc gagcaaacca tacttggggt
18721   gcaccattaa agacctccac ccctggggag cgatcctcac agcggagcag gacttggccc
18781   cttacaagtt tgcctagtct ctccctgaca ggtgggtaca gcagcaggag tctgaattcc
18841   acattttgca gagccctggc atagacggat cattcaagag aaacttcctt aagctgccca
18901   cggctgctca ggtcctcatg tctgcattca gatgtggcct tctttattac taactcagga
18961   ggaccttgat ggatgttagt ctccctggat ttgaaaacta tttatgggag cactgaaaat
19021   ggcattattc gttttgatgg caatcaagta cataaatgcc aatgatagag actgaaggat
19081   aagttccaca gcggatggg ttcaagtttc taaatactgc tgccatgatt ctcatgtgat
19141   gtgaaggata ttaagagatc tagaactagg atagaactgt agctcaacca tgctagaaaa
19201   aaatttttt aagttccact ctatttaaaa atatttgtgt agagacaaaa tctcactgtg
19261   ttgcccaggc tggtctcaaa ctcctggcct caagtgatcc tcctgctttg gcctcccaaa
19321   gtgccagaag ccactgtgcc aggccaaaaa gttatactcc tgagtaaaga ctactttata
19381   gtctcttgat aacctaccat aatttaaaga atattctgat tgctgtaaat ttctgtataa
19441   tttaaatctt aaggccgggg gcagtggctc atgcctgtaa tcctggcatt ttggggagcc
19501   aaggagggag gatcttctga ggtcaggagt tcgagaccag cctgaccaaa atggagaaac
19561   tccatctcta ctaaaaatac aaaattagct ggatgtggtg gcaggtgcct gtaatcccag
19621   ctactaagga ggctgaggca ggagaatcac ttgaacccgg gaggtggagg ttgtggtgag
19681   ctgagattgt gccattgcac tccagcctgg gcaacaagag tgaaactcca tctcaaaaaa
19741   aaaaccaaaa aatctttaaat ttgcagattt gtaaaatcct acatccctgt ccccaccctct
19801   ctttttcccat gatctcttagt ttagttcctc aaggtcagga atgagtctca tcttctctttg
19861   cacctctcac agtgcctgac acaatattgg ggacattgat aagtattgtt ttgtgtattta
19921   cttaggcata ttttctgtgg ttctctagca gtgggaaata aatgattgat gatcttatga
19981   agtcatagtc cacctaaaga ctgcgtagta cgcaaatgag tctaccatta tctcctcgat
```

88

```
20041    gtagattttc  atcttttatg  tatggtgact  ctaggtaagg  agaagaggtg  atctagctca
20101    cttggtggga  cgagcgtctt  gacacaaatc  cacaaaaaaa  ggagagtaca  cacagctgtt
20161    catttgtctg  tgtgttgagg  tggggtgggg  ggaggtgagg  ttgtattcac  acagagaaga
20221    catcttagct  ctgcacccaa  accaaaacga  gcgtcagtga  tgttaaagat  tgaaggtcag
20281    gtgctaccca  ggttttgtta  gtctcttgac  ataaacatgg  tgtctagata  tggtacatat
20341    tatgaatcat  tccatctttt  caaatacatt  ttgaagggtt  ttttttgtttg  tttgtttgtt
20401    tttgagacag  agtttcgctg  tgtcaccaag  gctggagcac  cgtggtgcag  tcttggctca
20461    ctgcaacctt  tgactcccgg  gttcaagcaa  ttctcgtgcc  tcagcctcct  gagtagcttg
20521    gattagaggt  gtgcaccatc  acacgcagca  aattttgta  ttttagtaga  dacggggttt
20581    caccatgttg  gccaggatgt  tgggattaca  ggcgtgagct  accatgccct  gccagaaggt
20641    ttgtttttat  tttaaacaca  attgtgctgg  gtgtgatagg  atgggtatat  gatagagcag
20701    aaaatcacta  gtatttcatc  aatatggata  gcatcctcca  taggttactg  atttaatatt
20761    gacaacaata  cccttcacac  gagagagagg  ttagagaatt  tgcccaaagt  caagcagctg
20821    ggaagaggaa  aactaggctg  tgctctccct  aaatctcatg  ctctattata  ttagggaaga
20881    ctctggcaga  tggcacatcc  tgattatttg  aaagtcacca  atatttttag  aaattgcata
20941    gataattaag  gttagtttat  ccagaattct  aataattgca  gcacatgtcc  ataaatctct
21001    attacaacca  cctgagatag  gccatctcgg  gcatgtggaa  aaagcccagg  tatagacact
21061    ttggatcctg  gatttccctt  gtactggctg  ctaccttggg  taaggtaatt  ctcctctctg
21121    aagctcctgt  cagcttctga  ggattgcgtg  ggatagcaca  taccaaatta  ctgcacacag
21181    agcctggtgt  acacttaagc  acacgagatg  tgtcaaagtt  ttcaaaaaca  ttgccaacga
21241    ggcatcagtt  acaaaacttg  ctgcagagtg  agctgatatt  gtgccactgc  actccagcct
21301    gggtgacaca  gtgaggcttt  gcctcaaaaa  aaaaaaaaaa  aaaaaagaa  agaaacacca
21361    aataagcaac  ttgctgcaga  aatgggtact  cttgttctag  aaatgtgact  atagggaagt
21421    tacaactacc  aactcgcgtt  aagggaaatg  agtgacctgc  cacctacatg  gtgttaggga
21481    ggttttgctg  agaaagtcac  tcatgaagaa  ggcaaaatat  agttaagaca  aaatgtaact
21541    atctatagag  ataaggtaaa  aattgaaat  agaacttcat  taaagatctt  tcaaataggg
21601    agaatgtggt  gaaaactgca  gttaacattt  gttaacagtg  tgatcatcgg  gttcagctta
21661    tcagtaacct  ggttcctgtc  tcttaactga  taaagaaaat  gggaggtttt  taaagagagg
21721    ctggctgttg  tatttagtaa  agctataaag  ctgtaagaga  aattggcttt  ctgagttgtg
21781    aaactgtggg  cagaaagttg  aggaagaaag  aactcaagta  caacccaatg  agggtaagtg
21841    gctttggggt  attttttcaaa  aatacttta  tctcaaaggg  aacaaaattt  tcacatacga
21901    atttaggtca  ttataacaat  aatattcatt  gttatattat  tataacatta  tatgttataa
21961    taatatataa  tatatataat  atataattgt  tataacatta  taacaattat  atattatata
22021    taatacaatt  atataatata  tattatataa  ttgtaatata  taatataatt  atataatata
22081    tattatataa  tataatatat  aatatatcat  atatgttata  tattttatta  tataatatat
22141    attatatata  atattatata  taatatatat  tatatataat  attatatata  atatatatta
22201    tatataatat  atttatatat  attatatata  atatatatta  tatattaaat  attatatata
22261    taatatatat  aacattattg  ttataatgtt  ataacattat  tatatattaa  atataatata
22321    tataatatta  tatttatatt  gaggaaaaca  ttcactcaat  gacagaggga  agacttagct
22381    ttccaaacaa  tcttctgtta  aaagaaaaac  tttagctgaa  ttaaatttaa  aggagtttaa
22441    ttgtgttttg  gtgccttatg  caaatgaagg  gcctataata  taatatataa  tgtaatatat
22501    ataacattat  aacaataata  atcattgtta  tcaattttgc  ttgtgagaaa  agtagtatta
22561    ccaatgggaa  aatcatggta  cagattgatg  aaatcccttg  aaatcttgtt  tgtgtaatat
22621    gggactttgg  tccaacgtgt  tggaaaagac  ctggttgagg  gtaagaaagg  atgggagctt
22681    tatagtatac  acacagagtg  ctcatattct  ggagttattc  tttctgaaat  ccacaaaatc
22741    cagcagtgta  atctactcag  atggagatga  aggaaaaaaa  caaaacaaat  gaagcaagaa
22801    gctgctggcc  aaaagggtg  aggaagcagg  gtcataccca  tctgcaagaa  gaaaatagag
22861    aacggacatg  agaggggtgg  atatgaggat  atgagaaggc  cctcttgagt  gagggccatt
22921    gtttctggag  gggtggggag  tgaatgactg  atgtaagttc  aagtttgttg  cgggtgcctg
22981    ctttgatcct  ctgcttggtg  gctccattct  tcttctctct  gaaggctact  tcaacagagg
23041    gtgcagaggt  aaacatagca  gaactggccc  ccaaagcttt  atccactgaa  tgtcagtttt
23101    tcctttaatt  tcttccaggc  ataaggaaat  cggaggcatg  tttctgtctg  gctataacag
23161    ttctcctact  agatcctgag  gccaaactat  tcacatgcca  ttgttcctga  ataccattgt
23221    ggacataatc  atctaatttt  ggctattcca  tgcagaggtt  tggagacaga  aagctctgga
23281    ctaatgcaga  tctgcattta  atccctgctc  ttacacttac  ttgttgcatg  accttaagca
23341    actcatctat  gtctttaaat  cttgagtttc  ttgcttacaa  agaggtgacc  agcatattag
23401    taataatatc  atgtttagag  ggtggtttct  gaaggttaaa  tgggtgtaag  tgtttagtgc
23461    attactcagc  cagcagtcaa  cacttaatga  tgtcagacct  catcattgtc  cttatttttt
23521    ttgttccaca  ctctgggaac  ttttatctct  gcgccctgta  gatatgctgc  tgtgactgct
23581    gacttgaatg  ttacgggatc  cccagcagga  tctgggattc  tctctgttct  tgaatgcgtt
23641    gtttaggtga  ggagattgat  tccaagttac  ttttctactt  gggtgcctct  gtgtttggag
23701    atgtggccgc  atagctcaag  tgtatctggg  gtgtcttcag  ttttcggata  cctgttttc
23761    ctctaggaga  attcttcctc  tatgtccttt  aaaatatttg  tcctataatc  acaagggttt
23821    atatttctct  gcccactcct  ccatacttcc  tggtttctg   ggagtgcaga  ggcagacaac
```

89

```
23881   tgcaagggag   agccccRcaa   ggatttgggt   gatgactttg   agccaggctg   acatccatgt
23941   ccaagggcaa   atgcagtctg   gttgggagga   tacaaatgcag   gagtcttaca   acagtggcga
24001   tgccaacatt   gtgggagcat   gcatgtgtgt   gtgcacgagt   gtgtgtgtgt   gcgtgcRctt
24061   gcatgcgcgt   gcatgtgcag   tagggtccta   ctacaggttg   gtggaaaaga   tccctctggt
24121   tctagaacaa   aaatctaaca   gcaaccaaga   gtccagttcc   aagaaagaga   aaggaaaggg
24181   ctagttaggc   agggcgtctg   gattctgaat   aggggctgag   gccttggctc   agacacagag
24241   gagaatggag   aaatggggtg   atgacagaga   catctttaaa   atatttgtgg   gccaagggta
24301   ggatggcttc   aacttgtctg   ggatgtattt   gtccttagca   cgcttggaat   aacatagcag
24361   cagcagcagc   agtcgacact   tactaagact   gctatgtgcc   agaccctgag   acaagcactt
24421   tccacacacc   acatcctgcc   attcttatga   tgatagtgct   gtgaagtagg   gatgaaacca
24481   agtccttgct   gactccaatc   ttgtacttcc   aaccaatgct   tgtgagggcc   ccaggccccc
24541   tacagcaaga   ctaaacatgg   acatatgtgt   ggctgggctg   agtctgtggg   attggctctg
24601   gagttgggtg   tgtctgagcc   ctgggtagaa   agggatgcta   tcctaatgaa   aaaacacaga
24661   ataaagatag   tgtccaccaa   gatgtaactg   agatttaaag   aagacttgat   aacttatcag
24721   ttgggtgggt   agtcttgaat   ttcaacaact   gcagatccta   tctaagacca   cttttgtctt
24781   gtgacaactt   agttaaacat   gactttagtg   tggatatcag   tggccaattt   tgagcccatt
24841   ctatgattgg   aacaaaggct   tccctgtcat   aggtgagtga   caactatctg   gtctggtggt
24901   acagggtaa   aagaatttac   taagacagtt   gtagataaag   aaaggcagat   ttattagaga
24961   aagtatgaaa   atacattaca   aggttgcaat   gggcagcaca   gccagagagg   agctgactgc
25021   aaagaaacaa   aggcttgctg   gatatttata   ggatagttcc   tgggctccag   ggggctatat
25081   gcagtactga   taatgccaag   gttgtagtga   gctaacttgc   aggtgtctgg   tgatagttgg
25141   gtgcYggaag   attctgagtt   atttgctcag   aagagctgtg   tgtcctgggc   catgaagaaa
25201   ggcaaactta   cagtttatct   gctttctgtc   tttgctttcc   cttggtactg   ctagtctgac
25261   ttttttttcc   ctaattagga   ttccgcacat   cctatgcctK   tctctcagga   gccttgcagt
25321   caaagtgtga   tggagtgtct   gagggccgag   tcattgagta   atggggcctc   tgccctctca
25381   gccccactgc   cacccaactg   aggcttcctg   gatgagggtg   tatgccgtgg   ggaaaaggag
25441   tgggatgatg   tggatccccc   tggggttgca   ggtcatcctc   ccaaatcccc   actctccctg
25501   tctacccaca   ggaccagtca   gcatttctct   aattcctctt   cactctgggg   cattatgatc
25561   agaaaaatatc   atgtacctt   tgctgcaatc   ttggctgaaa   gaacctcaca   ctatacttga
25621   atgtctttga   tttaccatct   gagcagagat   ctcaaactat   aaagaaaatc   agagtctctt
25681   gatgcccatg   ttcaggacaa   gtttcttgg   atttcatgac   tctttaggaa   cagtgagctt
25741   atgggtaaag   accgagactt   tagtgaggaa   ttccctcatc   attttttctct   tcttgagcaa
25801   aaccttgtct   gatttggggt   tgtaatgtgc   aaaattattt   tcttttcatt   tcacctttt
25861   ttcttaagta   gtggtttctg   gaaatctgag   ctgctttgtg   caccctgact   agtctgtgcc
25921   agatcacgaa   cgtacacacc   aactctcaag   acaccagaaa   aatcacccтt   tagaggaatt
25981   atctagaaaa   ctagcccta   tgcaaagtga   ataacctttt   caaaatataa   tctacattaa
26041   gacttttgaa   actaacacat   tcctcaatca   tagcaaccaa   atgaattact   taaagtgggt
26101   ttggttcctg   ctttcattaa   gggagtattt   tcaggaaatt   agaggtctaa   tgtaaacatg
26161   gatcctttag   gttttttccat   tttctttgtc   ttcaggctgt   ttgctgccac   tttcatcctc
26221   cccttatctc   agtctgccag   tgatggtttg   tgcctctggt   cКctgggtca   gatccaggac
26281   taattcactc   acctctgatt   tctagttcca   cMcttatgac   gaaagcattc   ttaaatctga
26341   tattcacata   aaaaaggatt   tgattgtctt   tggcatgtct   ctgcagttag   ttaatcatgt
26401   cttgagcaca   gaaaaatggg   aagaaaatgg   aaataagatt   tctttgatag   gaggagtcct
26461   gctaagactt   gtctgtccct   gagctgagat   atgctaattt   accttaaatc   tttccataaa
26521   ccacattctg   aattccagtt   tataaatctt   tatcttattt   cccgaacctc   atctcatact
26581   ttttatgtRa   tctggtgggc   atagatattc   tcttggtgaa   cctgaggaaa   tctagaatga
26641   caaaggatct   aaaataagta   acacagttgc   tcaagttcct   gggccctatg   gtagaatttc
26701   cttcttcaga   gaaggggaga   acaaggtaga   taaagaaagg   agcatttcat   cttgcagcag
26761   cttggaagat   agctgggagt   ttaattttta   taaaaagaa   agcaatcaaa   ccatccaata
26821   cagcccattt   tcaaccaaaa   gtatttcaga   attacctgga   gatttatcta   aaattatgta
26881   aaaattacct   agaattacct   aaaaaaattg   tacaatttca   tttcccttct   gttacatgat
26941   atggataccc   agctactatt   tattagtagt   agatacagac   tatttaagag   acaaaatatt
27001   atattccaaa   aagtgtgacc   cttagaaacc   aaaattccaa   agaagagttc   ttgtaattgg
27061   caaggctggc   tctgattgag   tgtcaaactt   cctgtacaat   gatggatcaa   cttagaaatt
27121   tgatagattc   aaagcattct   gaaattgtac   aactcgctca   gcacttgtat   tatttacatg
27181   tcttctgtta   ttttccattg   aatgatgttt   tcatgtatac   tcagctgctt   ataagggcaa
27241   tattttaaaa   tcagagcaat   ttctttaaaa   ttaaaatgtt   tcagcccagc   atggtggctc
27301   aYgcctgtaa   ttccagcact   ttgggaggcc   gaggcaggtg   gatcatgagg   tcaggagWtc
27361   gagaccatcc   tggctaacac   ggtgaaaccc   catctctagt   aaaaatacaa   aaaattaact
27421   gggcgtggtg   gtgggcgcct   gtagtcccag   Stactcggga   ggctgaggca   ggagaatggc
27481   gtgaacccag   gaggtggagc   ttgcagtaag   cccagatgcg   gccaccgcac   tccagcctgg
27541   gtgacagagc   cRgactctgt   ctcaaaaaaa   aaagtttcag   caaaatactt   tttaaaatgc
27601   actgtaatta   ccctaacaat   attacactgt   tacagatagt   ttaagaataa   taaaaaatgt
27661   aattcataaa   ctcactcctt   taatacaaca   aacctgatta   atgttctgtg   ttccttatta
```

90

```
27721    gactttctcc catgtatagc tttacaaagt ggtagttgat gagtaagtac aatttgtgat
27781    atttttttca cgtgagaaca aactcaatat gtcaattact tcatagattt ttaatttata
27841    aaaattttct catatctctt ataaaaataa aaaaagattt gatttcagta aacacacatt
27901    tggagagaaa gttgatcctc atataaactc aaccatagac cattcagtgt gttgcaagtt
27961    ataaatgcca tggaggaaaa aaatcaagca gagttaggaa gattgggagc acagggttag
28021    ttggtcaaag ggtgggagag gatagacctc attaagaaag tggcagacct cattgagaaa
28081    gtgagggagt gagtcacaca gacacctgga ggaagaaaat gcaagagagg gtaaacagtt
28141    gctgcaacac cctgaggtgg gagcaggctc gggaggccaa gcaacaggaa ggacaaaagt
28201    gcacgtgttg tgtagttaag attggggggtg tagcaggatg taaatcagag aggtgagagg
28261    tgacagaccc ccagggggtc ttcttgttca ttgcaagcaa tttggctttt gttgggtaaa
28321    gcagaggttt gaagagagga gtaatttgaa aggatcactc tagcttctgc tttgaaaatc
28381    aattataggg gacaagggaa ggatcagtgc aaagactagt gcaataattc aggcaagaga
28441    tggtggtggc ttggatcagc aaagatggcc atgagaagtg gtcacatggc tgagttctgg
28501    ctgtatttgg aaagtaggtt cactgggact tattggtggc ttgagattga tgtgctggag
28561    aaagagaggt ataaggatac ctccaaggat tttgatatgt gcagctgaaa aaagatgaag
28621    ttgccatcaa ccaagatgag aatgatggtt ggaagagcag attttgggag aagaacaggg
28681    attcattctt gacaggttat ttttgtgatg tctggaatat agacaagtga aggcattgag
28741    tcagcaggtg gacatgagaa cctgaagttc aggggaaaag tgtggacggc aggtgtaact
28801    gcaagagtgt aatcagtata gagatggtgt ttcaaaccag ttaactggat gacatcatgt
28861    ggggagtgag tgtggacaga gaagaaaagt gcttcaagga ctgaaacttg aggtggtgga
28921    gagaagagga gatgtgaagc agtagcctgt gggtcctgaa gaaagatgtt ggagaaattg
28981    ttccactttg acagaggctg cttatagctc aaggtaaggg gaagacgcag gactgaggat
29041    tcctaagcaa catggtggta attgacgata ttgataacag cataaagaca tcataattta
29101    aactaattat taacaaaatt gaaatttaaa gaaacaagta atgcaatgtt tccaaagtgg
29161    attcaaattc cataatatta ttgcaaaaca aacaaaccaa acaaacaagg aaaacaaaga
29221    aagaagggat tctaacataa aagaatttaa aaataagtaa atcagctcgg aaaactcagc
29281    aggctatatc ttcctggtgg ttagaaatat ctgcttatta aaaggtttca aaggtccttt
29341    gggaatagaa cttattgcaa tgtgtttggc tgtacatttc ttacactttt tttttttaact
29401    gaggagtcaa ttaaaatatc ctgaagtgac tgtccataga acacattttg aaaacacaaa
29461    tgctatataa tcacaaatat atctccctcc cctttgctag tgattgttac caaactgcct
29521    ctgctagcta ggattcaact ctttctctta acaactatgc tttgatggaa caattttttt
29581    agatacaaat attgtcaatg agaacaatta ttatacagtg cagtcacaga agagaaacag
29641    aaaagaaagg gaggtgcaac aaaatgaagg taaatattag ccttcacact taaatacatt
29701    tgatgagggg ataaaatgta catgtaggtg tggatttatg ggctgtgttt ttcaaacaat
29761    taccaaatat ttattaacca ccaactcagt gtaagctagt acactaggtg cctatagtga
29821    tccagcaggc tataaggcag gatttctgcc cttatagaat tttggggcat attctcagta
29881    tcttttgtat ctctcgccct cttttctttt ttttgtgaca caagatgtca ttctgtcatc
29941    caggctgtag tgcagtgaca tgatcatggc tcattgcagc cttgacctcc caggctcaag
30001    cgatcctctt gcctcagcct cccaagaagc tgggaccatg ggcatgtgcc atcatggcag
30061    actaattttt aaatttttg tagagatggg gacttactat gttgcccagg ctagtgttga
30121    actcctgggc tccagcagtc ctcccacctc agcctcccaa agtgctggga ttactggtgt
30181    gaaccatcac tcccagccct ttcttttgct ttgttatttt aaaaactttt ctttattctc
30241    tttctgttac ctttttaataa actataagaa gtctttcaaa acaaaaatgt cttaatgtca
30301    gtgggttagg ttttgtcctg cattattttc ttttaccoca tggtggggaa aaaagcaaaa
30361    aactgtttgg tgcctataca atgccttatc gcacagcttt ggtttgtgtt tactccagtg
30421    tcgtctttaa ggtggtgagg tcagtgcagc tggcctgggt cagaactcag ggtcacttca
30481    tgccagatgc tcaccacctc tgcagtgttg tatgcagaga tccaaattca tccttcaggg
30541    aaggatgaac acttattctc tgggatgttt tacagtgaaa atcagaagtt gctctattgt
30601    gccctcatgg cctagctact tccactcttt tgaagatctt acctggtctc atgggtttaa
30661    tttcatctat acagtgacat atcccaaatt tgtatctcta gcccaatatg tctcctgaat
30721    ttcagaagca Ratattcatc ctactgatgg gttcacatct ccagttttat gtctaatgag
30781    gtgtcaagct gaatatggcc aaaactagtg taatcctact ctaaaaccta taccacctgc
30841    catttcctaa ttcagtaact agcaaccctta tttgaccagt tactcagggc cccaaacctc
30901    tcttgactgc tctcatatcc tgtccatcaa gaagtcctac tagttctacc ctcaaaatct
30961    atctgcatcc caatcatgtc tcaccacctg cacttccccg ccctggtcat agccaccgtc
31021    tctgtctcac ctggatacct catcccttgc cattcttaag agtgttttca cattcagagt
31081    gtttaatacc tacgtcaagt ttgaggtcat gtcactcctt agcttgaaac actccagtgg
31141    tcttcatcac acttaatada agccaaggtg ctcagggtga ctcactcgca atctagtctc
31201    taccctctc caaccagctc tcactgctct ggagttcctc atgccctgtt ggccatgtgg
31261    cttccctact ggccctcgga cacaccaggc ttgcctccat ggctggactt tcatgctgcc
31321    tgccctttcg acctgggctg tttattaaat gtctcctctcc tagttggcct tcccaggaca
31381    tctggcctga aattgctctc ctcctgctgt atttttctcc ttaagtggat cacactctaa
31441    gatcttttgt aactttttttt tatttctttta tgttacttat tgtcctcctg taggacatga
31501    attccacaag ggcaggagat ttttgtttcc ctttgctctg tgtggtgcat tgctgtatcc
```

91

```
31561   ccagcatctg aaacagtgat tggcatgaaa gaataaccct cactggttgt ctgggttcag
31621   tattttgctt ttctaagctc agttgatgca gaacaactgg tttttctgggt tcagtatttg
31681   cttttctaag cccagttgat gcagtgttgg tatgtgcggc catgttagat gttagaacac
31741   tgacatcctt accactgatg agtcaatagc cactgccccc gaccctcctg atattttatg
31801   gtccgagtaa ttactagatg catctacatt taaggctgca aacatgaatt tactggagga
31861   agcacagtgg tggtggatga cccaatctct ctctcaaact taccaaaatc accaaggaat
31921   cataaatcca cccctctttg catcttccaa gcaaaggaaa ttgtagaaaa cacaagtgaa
31981   actgtttcac agagatatca gcaatacata gcaaatttga gatgatggag attaagagga
32041   aaattaaatc attcttaagt ggtcaaggta aaaacattat tgtgYggttc gctatagtta
32101   catttaaaac agtccgaagt caaacaactc tagtctgaYt agtcttgtca ttgatatgtg
32161   ctgtcttaga attctccaga gagacagagc taataggaga ttatctatct atgtatctat
32221   gtatctacct aatctatcta tactctctct atctattta aggaactgac tcatgagatt
32281   gtggggcctg gcagtctaaa agctacaagg caggctggca gactggaaat tcagataaga
32341   gttggtgttg tagtcttgac tctgaattcc ataaggcagc cagtcggaat ttcaggcggg
32401   gtttctgtgt tgcagtcttg aggcagaatt ccttcttctt tgggaaacct cagtctttgc
32461   tcttcaggcc ttctgctgct cacattccac agtgctgtcc ttctggccat ccttcatgct
32521   tttgtgttgc ctgtcccatt gcccttaatc ttatatggaa aaatggaaat aaattcagag
32581   cggaaactta tctaaaagtt ttgagggaaa aatgatcgaa accagaatta ttcacaggac
32641   aacttgttct atgaataatt tttgaaaagg aaccatttcc tctacagctg gcttattatg
32701   aggagcttga tgataaaata ttggcaaggc ctccagtgga catttatcgt gttgtgacag
32761   atcccagaaa tgaatctggg tctcccgact tcaatacagt tctcttctta ttacataatg
32821   agagaatatg attttctact cttttctttt actttacttg ttgagaaaca actaatattc
32881   aattaaaagc atcaatctta catattcagt tgaatgagat tcagcaatca tatcccccat
32941   gtatcatcat ccatatcacc atcccaacag cacagaaagc atttcttctt cccagagagt
33001   tgccctcaca ccttttccaat caaatttcca catactctct agtttccagc atgatagatt
33061   ataatttgcc tattgttggg cttcctatag gtatgatatg taatatgtag tcttctgtgt
33121   ttagtttctt ttgcttagca taatttttca ggagcaaggc tttattattt ttttgtttgg
33181   aaaYgaactc agagttacat aacatgatgt ttttgacagt catcctatgg ttatatcagt
33241   atttccttcc tttttcattgc tgagtagttt acattaaaaa aaggatccat tctcctgttg
33301   atagacattt ggattatgtt taattttgg tattatgagt atggcattta tgaacattct
33361   tttacaagtg ttttgtggat atgtttgtta attctcttgg gtaaggggg agttattgag
33421   ccatacagta gatgttcttt aacttttgt taggttttat atatatat atatatat
33481   atagtgtgtg tgtgtgtgtg tgtgtatata tatatggtgt atataattta tatacactaa
33541   aatacacaca tttttaagtg taaagttcaa taagttttga tacatgtata ctccattggc
33601   tttgttttgt attgctataa aataatacct gaggctgggt aatttattaa gaaaaagttt
33661   atttggctca tgattttaat ggctgcaaag ttcaagactg ggcaactgca tctggtgagg
33721   gcctcaggct gcttccaatc ctggtggaag gttaagggga ggtggcttgt gcagagatca
33781   catgggagag aagaagcaag gggtggggga gtgctaagtt cttttaaca gccagctccc
33841   atgggaagta atagagtgag aactcaccac caagggaggg gattgaagga gggaattaat
33901   ctattcatga gggatccacc cccatgaccc aaatgctccc ccttaagccc caccttcaat
33961   attggggatc aaatttcaac atgaggtgag gagggtacaa atatccatgc ataccacctc
34021   tgaaaccatc atctcaattg acatattcca caaatctctc tcatactcct ttataagtaa
34081   ttgctaccct tcatcttcca ttctaggtac ccactcactg gccttgcagt ttgggaacat
34141   tatataaatg ggacgatata atatgcattc tttttaaaact ggcttctttc attcagtatg
34201   tttctaagat tcatcagtag ttctttttt tttttttgct gagtagtgtt ccattgtatg
34261   aatatagaat agtttagcca tatatgtgct gacggtcatt ggtcttgttt ccagtatttg
34321   gctatatgaa ttgaactggt atgaacattt gtgtacaagt cttttgtgt gcataagttt
34381   ttatttctct taggtaaatc cccaggaaga gaatttctgg agtgtaacac aagcacataa
34441   tcagaaactg ccaaactatg ttcactgtat catgttacac tccatcaatg atgcatgaga
34501   gttctaggtg tgccacatcg ttgccaaaat ttggtattgt cagtctttt tattagctat
34561   gaaatagaat ctcgatattg tggagccaag atggccaaat aggaacagct acagtctaca
34621   gctcccagcg tgagcgacac agaagacagg tgatttctgc atttccaact gaggtactgg
34681   gttcatctca ctggggcgtg tcggatagtg ggtgcaggac agtgggtgca gagcactgag
34741   cgtgagccaa agcagggcga ggcatcgctt cacccaggaa gtgcaagggg tcagggaatt
34801   ccctttccta gtcaaagaaa ggggtgacag atggcacctg gaaaatcggg tcactcccac
34861   cctaatactg tgcttttcca atggtcttag caaatggcac accaggagat tatatcctgc
34921   acctggcttg gagggtccta cacccgtgga gcctcgctcg ttgctagcag agcagtctga
34981   gatcaaactg caagttggca gcgaggctgt gggaggcgtg cccaccattg ctgaggcttg
35041   agtgggtaaa caaagcagcc gggaagctcg aactgggtag agcccactgc agctcaagga
35101   ggcctgcctg cctctgtaga ctgcacctct gggggcaggg catagctaaa caaaaggcag
35161   cagaaacctc tgcagactta aatgaacctg tctgacagct ttgaagagag tagtggttct
35221   cccagcacgc agctggagat ctgagaacgt acagagtcac tcctcaagta ggtccctgac
35281   cccgagtag cctaactggg aggcacccc cagtagggg cagtctgaca cctcacacag
35341   ctgggtactc ctctgagaca aaacttccag aggaacgatc aggcagcaac atttgctgtt
```

```
35401   caccaatatc cgctgttctg cagcctctgc tgctgatacc caggcaaaca gggtctggag
35461   tggaactcca gcaaactcca acagatctgc agctgagggt cctgacagtt agaaggaaaa
35521   ctaacaaaca gaaacgacat ccacaccaaa accccatctg tatgtcatca tcatcaaaga
35581   ccaaaggtag ataaaaccac aaagatgggg aaaaaacaga gcagaaaaac tggaaactct
35641   aaaaatcaga gcacccctcc tcctccagag gaacacagct cctcaccagc aatggaacaa
35701   agctggatgg agaatgactt tgacgagttg agagaagaag gcttcagatg atcaaactac
35761   accgagctaa aggaggaagt tcgaacccat ggcaaagaag ttaaaaacct tgaaaaaaaa
35821   ttagatgaat ggctaactag aataaccaat gcagagaagt ccttaaagga cctgatggag
35881   ctgaaaacca tggcacgaga actacgtgat gaatgcgcaa gcctcagtag ccgatttgat
35941   caactggaag aaagggtatc agtgacggaa gatcaaatga atgaaacgaa gtgagaagag
36001   aagtttagag aaagtagaat aaaaggaaac gaacaaagtc tccaagaaat atgggactat
36061   gtgaaaagac caaatctaca tctgattggt atacctgaaa gtgacaggta gaatggaacc
36121   aagttggaaa acactctgca gggtattatc caggagaact tccccaatct agcaaggcag
36181   gccaacattc aaattcagga aatacagaga acaccacaaa gatactcctt gagaagagca
36241   actccaagac acataattgt cagattcacc aaagttgaaa tgaaggcaaa aatgttaagg
36301   gcagccagag agaaaggtcg ggttacccac aaggggaagc ccatcagact aaaagctgat
36361   ctcttggcag aaactctaca agccagaaga gagtgggggc caatattcaa cattcttaaa
36421   gaaaagaatt ttcaacccag aatttcatat ctagccaaac taagcttcat aagtgaagga
36481   gaaataaaat actttacaga caagcaaatg ctgagagatt ttgtcaccac caggcctgcc
36541   ctacaagagc tcctgaagga agcactaaac atggaaagga acaaccggta ccagccactg
36601   caaaaacatg ccaaattgta aagaccatca aggctaggaa gaaaatgcat caactaacga
36661   gcaaaataaa tagcaaacat cataatgata ggatcaaatt cacacataac aatattaacc
36721   ttaaatgtaa atgggctaag tgctccaatt aaaagacaca gactggcaaa ttggataaag
36781   agtcaagacc catcagtgtg ttatattcag gaaacccatc tcacatgccg ggacacacat
36841   agactcaaaa taaagggatg'gcggaagatc taccaagcaa atggaaaaca aaaaaaaggc
36901   aggggttgca atcctagtct ctgataaaac agactttaaa ccaacaaaga tcaaaagaga
36961   cacagaaggc cattaataa tggtaaaggg atcaattcaa caagaagac taactatcct
37021   aaatatatat gcacccaata caggagcacc cagattcata aagcaagtcc ttagagatct
37081   acaaagagac ttagactccc acacaataat aatgggagac tttaacaccc cactggaaac
37141   attagacaga tcaacgagac agaaagttaa caaggatatc caggagttga actcagctct
37201   gcaccaagca gacctaatag acatctacag aactctccac ccaaaatcaa cagaatatac
37261   attttttca gcaccacacc acacctattc caaaattgac cacatagttg gaaataaagc
37321   actcctcagc aaatgtaaaa gaacagaaat tataacaaac tctctcttag accacagtgc
37381   aatcaaacta gaactcagga ttaagaaact tactcaaaac cgctcaacta catggaaact
37441   gaacaacctg ctcctgaatg actactgggt acataacgaa atgaaggcag aaataaagat
37501   gttctttgaa accaacgaga acaaagacac aacataccag aatctctggg acatatttaa
37561   agcagtgtgt acagggaaat ttatagcact aaaaagcccac aagagaaagc aggaaagatc
37621   taaaattgac acactaacat cacaattaaa agaactagaa aagcaagagc aaacacattc
37681   aaaagttagc agaaggcaag aaataactaa gatcagagca gaactggagg aaatagagac
37741   acaaaaaacc cttcaaaaaa tcaatgaata caggagctgg ttttttgaaa agatcaacaa
37801   aattgataga ccgctagcaa gactaataaa gaagaaaagg gagaagaatc aaatagacgc
37861   aataaaaaat gacaaaggcg atatcaccac cgatcccaca gaaatacaaa ctaccatcag
37921   agaatactat aaacacctct atgcaaataa actagaaaat ctacaagaaa tggataaatt
37981   cctgttcaca tacactctcc caagactaaa ccaggaagaa cttgaatctc tgaatagacc
38041   aataacaggc tctgaaattg aggcaataat taatagctta ccaaccaaaa aaagtccagg
38101   accagatgga ttcacagccg aattctacca gaggtacaag gaggagctgt taccattcct
38161   tctgaaacta ttccaatcaa tagaaaaaga gggaatcctc cctaacgcat tttatgaggc
38221   cagcatcatc ctgataccaa agcctggcag agacacaaca aaaaaagaga attttagacc
38281   aatatccctg atgaacattg atgcaaaaat cctcaataaa atactggcaa accaaatcca
38341   gcagcacatc aaaaagctta tccaccatga tcaagtgggc ttcatctctg ggatgcaagg
38401   ctggttcaac atacacaaat cgataaacat aatccagcat ataaacagaa ccaatgacaa
38461   aaaccacatg attatctcaa tagatgcaga aaaggccttt gacaaaattc aacaaccctt
38521   catgctaaaa actctcaata aattaggtat tgatgggacg tatctcaaaa taataagagc
38581   tatctatgac aaacccacag ccaatatcat actgaatggg caaaaactgg aagcattccc
38641   tttgaaaact ggcatgagac agggatgccc tctctcacca ctcctattca acatactgtt
38701   ggaagttctg gccagggcaa tcaggcagga gaaggaaata aagggtattc aattaggaaa
38761   agaggaagtc aaaattgtcc tgtttgcaga tgacatgatt gtatatctag aaaacccccat
38821   catctcctta agctgatagg caacttcagc aaagtctcag gataaaaaat caacgtgcaa
38881   aaatcgcaag cattcttata caccaataac agacaaacag ccaaatcatg agtgaactcc
38941   cattcacaat tgcttcaaga gaatgaaata cctaggaatc caacttacaa gggatgtgaa
39001   ggacctcttc aaggagaact gcaaaccact gctcaatgaa ataaaggagg atacaaacaa
39061   atggaagaac attccatgct cataggtagg aagactcaat atcatgaaaa tggccatact
39121   gcccaagtta atttatagat tcaatgccat ccccatcaag ctaccaatga ctttcttcac
39181   agaattggaa aaaactactt taaagttcac acagaaccaa aaaagagccc acattgccaa
```

93

```
39241   gtcaatccta agtcaaaaga acaaagctgg aggcatcacg ctacctgact tcaaactata
39301   ctacaaggct acagtaacca aaacagcatg gtacttgtac caaaacagag atatagatca
39361   atggaacaga acagagccct cagaaat:aat gctgcatgtc tacaagtatc tgatctttga
39421   caaacctgac aaaaacaaga aatggggaaa ggattcccta tttaacaaag ggtgctggga
39481   tatctggcta gccatatgta gaaagctgaa actggatccc ttccttacac cttatacaaa
39541   aattaattca agatggatta aagactgaaa ttttagacct aaaaccataa aaaccctaga
39601   agaaaaccta ggcaatacca ttcaggacat aggcatgggc aaggacttca tgtctaaaac
39661   accaaaagca atggcaacaa aagacaaagt tgacaaatgg gatctaatta aactaaagag
39721   cttctgcaca gtgaaagaaa ctaccatcag agtgaacagg caacttacag aatgggagaa
39781   aatttttgca atctacttat ctgacaaagg ctaatatcca gaatctacaa ttaactcaaa
39841   caaatttaca agaaaaaaac aaacaacccc atcaaaaagt gggcaaagga tatgaacaga
39901   cacttctcaa aataagacat ttatgcagcc aaaagacacg tgaaaaattg ctcatcatca
39961   ctggccatca gagaaatgca aatcaaaacc acaatgagat accatctcac accagttaga
40021   atggtgatta ttaaaaagcc aggaaacaac aggtgctgga gaggatgtgg agaaatagga
40081   acacttttac accgttggtg ggactgtaaa ctagttcaac cattgtggaa gtcagtgtgg
40141   cgattcctca gggatctaga actagaatta ccatttgact cagccatccc attactgggt
40201   atatacccaa aggattataa atcatgctgc tataaagaca catgcatata tatgtttata
40261   gcggcactat tcacaatagc aaagacttgg aaccaagcca aatgtccaac aatgatagac
40321   tggattaaga aaatgtggta catgtacacc atggaatact atgcagccat aaaaagatga
40381   tgagttcatg tcctttgtag ggacatggat gaaactggaa accatcattc tcagcaaact
40441   atcgcaagga caaaaaaccg aacaccacat gttctcactc atagatggga attggacaat
40501   gagaacacat ggacacagga aggggaacat cacacaccac ggcctgttgt ggggtggggg
40561   gaggggggag gcatagcatt aggagatata cctaatccta aatgacaagt taatgggtgc
40621   aatacacaaa catggcacat gtatacatat gtaaaaaacc tgcacgttgt gcacatgtac
40681   cctaaaactt aaagtataat aaaaagaaga agtagaatct cattgtgatt ttaattgttt
40741   acatttatct aataactagt aactttgtgc actttgtcct gtgcttattg aacattgata
40801   tattttcttt tgtgaaatat cagtacaagt cttttcctca ttttaaaaat tgggctattt
40861   tatctttttc atatgaactt gtaagagtta tttatgaagt atagatacaa gtctttcatc
40921   tgatatgcaa attatgagta ttttgtctgg actgtggctt agtctattca aattctaatg
40981   atgtctttga agaagaaaag tttctaattt taataaagtc ccagctatga tttattttc
41041   tttcatgatt agtgattttt gtatccattt aaaaaatatg ttcctacttc aggtttaaga
41101   tgttctgtta tgtaatacaa tccatttttа tttcatttttt ctgtataata tgaggtagag
41161   cttttttcaca tagagagtca atttttctat cattatgtgt cgaaggcagt gttttttcaca
41221   ataaacttat gttggcactt taatcaaaca tcaatttact atatattttt gagtctactt
41281   ctggactctc tcttctttga tcaacatgtc tgtctttatg cccataccac actgttttca
41341   ctatggtagc ttttatatta agtcttaagg tcaggttgtg taagttctct gactttgttc
41401   tttttctttta accatgtctt gggtattccc aggactttaa tttgccatgc agactttaga
41461   tttccataga agctttagaa gcagattctg ctgctgggag aataagttac actgaccca
41521   ggacaaaaga tgtacatgaa ctaaataaag gacgacaaat gcagaccaca gcaataaagt
41581   ctgtggaatt agtatgtgag aaaattacaca atggtctaaa atattagtct tgaagggatc
41641   ttaatgatga tttagtcatt tcacatgtgc aaactgttaa gactgctggc tgttttttcac
41701   atggccacca gcccatttct ttttcctgag tgtctgcagt ccacagagct tcttctggct
41761   ctctatgctc catctgttct cggcccagga gttcccatgg gacctcccaa gggtgactca
41821   ggtctcatct tccgataaga taaaagaaat aagacagata ttcgggggtag agatgtaaga
41881   gcatggtgaa agacagactg ctcctgttta gatccatata taaaagcttt attttttgaa
41941   tcccaaattc catcccttgg tcatctgtac tggttcattc aataaatatt tctgaggctg
42001   ggcatggtga ctcaagcctg taatcccagc actttgggag gtcaaggttg gtggatctct
42061   tgagcccaga tgtttgagac cagctttggc aatatggtga aaccccatct ctacgaaaaa
42121   tacaaaaatt agctgggcat ggtgatgtgt gcatgtagtc ccagctattc tggaggctga
42181   ggtgggagga tcacctgagc ctgggaggtg gagactgcag tgagccatga tcacaccact
42241   gcactccatc ctgggtgaaa gagtgagacc ttgcctcaga ataataata ataatgataa
42301   taagtaaaca tttctgagca cctacttggt gctcttccag gcacatgagc tatatcaata
42361   gacaaaactc acaaaaataa tMttgccttc tggaggtcat attccaatgg ataaaacaga
42421   aaataataaa catgatgcat aagtagaatg ttagaagtga tgaatgcctt agaaaaagat
42481   tggagcaggt taaggagatg gggattggga gaggagacat tgcaattcaa aattgattgg
42541   tcaggaactc aaacaaatac ttggacacYg atagcagcat tattcaaaat agtcaatagc
42601   tagaagcaac ccagatgccc attaagggat gaatggataa acaaagtgtg gtaggcacac
42661   acagtggaat attattcagt cattagaaac aataaagtac agacccatgc tacaatatga
42721   acaaaacttg gaaagattat gctaagtgaa ataaggcagt cacaaagggc acttattgtg
42781   tgactccatt tataggaaat attcagaata gataaaccca tagaatcaga aagcagattg
42841   gtggtggcta tgggtttgtg gtggaagagg aagaatgggg acttcatgtt aatgggtatg
42901   gggttatact cggggatggt gaaaatgttt tgaaactaga tagagttagt ggttgtacca
42961   cattgtgaat gtaccaaatg tcactaaatt gtgcacttcg aaatagttaa ttttatgtta
43021   tgtaaatttc agcgcaattt ttaaaaatca gaagcaaaaa aagaacagaa aattgagtag
```

94

```
43081    ttaggctaag  tctcacagag  aatgtacagt  ttgagtgctc  acttggagga  ggtaagttag
43141    ctccttggagg aggtgggtaa  caggaagaag  ctgtccaggc  agaaggaaca  atgtctgcct
43201    agtgccttca  aggaacagca  atgaggatac  ggggttggag  gggaggggtc  agagaggtga
43261    ttggtggtgt  tccatcgtgt  ggtacctccc  agcttctgta  aagacgttgg  cttttcctct
43321    gggtgaggtg  gaagactttg  gaggccccct  tgatttgacc  taggttctaa  cagaatctcc
43381    agagtgtctg  ttttaggaat  atacttaaga  gggtacaagt  ggaagcaggg  atagcaattt
43441    agaagcagtt  attaaaatcc  aagtgagaga  ccatggtggc  ttggacaagg  atggtggcat
43501    tgagaatgga  gagatgcaat  tgcaggtagg  atgtcctttg  aagatgaaac  cataagttgt
43561    cctgatggat  tgcatgttca  atatgtgagg  aagagaagag  tcaataataa  atggagtctt
43621    ggtctgagcc  aacattcagc  atggagttgt  catcattaca  attagaagac  aatgggagat
43681    gctgaatttt  ctttggggca  agggaggcaa  ggaggctgga  ggagatcaga  agttctgttc
43741    tggagacatg  aagcttttgt  gaattccact  tgttccttga  atatatgatg  gtttcttctg
43801    ttttgtttg   tttgttttt   taaattttag  attctaaggt  acatgtactt  gtttgttaca
43861    aaggtatatt  gcattactgg  tggggactgg  tcttacccaa  attgtgaata  ttgtacccag
43921    taggtaattt  ttcattcatc  accccctacc  ctccctactt  ttggagtccc  cagtgtctat
43981    tatttccatc  tttatgtcca  tgagtaccca  tggtttagct  cccacatatg  agtaagaaca
44041    tgtggtattt  ggttttctgt  ttcttaactg  gttcacttaa  gataatggtc  tccagctcca
44101    tccatgtagc  aaaggacata  atctcattct  ttttatggct  gaatagtaat  ccgtggtgca
44161    taggtacaac  attttcttta  tctagtcaac  cgttgatgga  cacttgagtt  ggttccatga
44221    ctttcctatt  gtgaaccgtg  ctgcagtgaa  catactagta  gaggtggaga  ggtgtctttt
44281    ttatacacta  atttctttc   cttagaaaag  tagtgagatt  tctgggttga  atggtagttt
44341    tacttttggt  tcattgagaa  atcttcatac  tgttttcctt  agagattgaa  ctaattgaca
44401    ttcccagcaa  caatgtataa  gcattccctt  ttctccatat  ccatgccaac  atctgctgtt
44461    ttttgagttt  ttgataatag  ccattcagac  gggtcttctg  gttttggag   gggaggattg
44521    ggtgaagcaa  gaaggagttt  ggagggaagg  agtggaggct  tgagtgggcc  tagagtttgg
44581    agtatgggca  agaaggatcc  cagagacaag  cactttgccc  acagctacac  agctaatgga
44641    gctgggggc   ccagcatatt  ctccagaggc  ccaggctagg  ccgtctaggg  gcatgcttgt
44701    tctgccccat  ccactgcagg  tcctgaaaac  attttcatca  ataaaaaaaa  ataaaaaaac
44761    agaataaaaa  tgatacccag  atgtcctcta  gtgaaatgag  gggaaaaaaa  acatccatcc
44821    ccRgcttatt  gtgagagtca  ctgaatgaga  gcctgactca  ctgaaattca  tataagatta
44881    atgtaaccaa  gttccctgtc  ttttgacact  ggtttacagt  aagagcaggc  cacacatggc
44941    cagctctgga  gtgtgttagg  acatttacat  tttacatttc  agtgtgatat  ctgctaagtc
45001    aaatgaagaa  gtcttgaaag  ataccttcta  agttcggaag  tatttgagtg  tcacattcca
45061    tatgccagca  gttagttgtt  gccctcaaaa  cataaggttt  tgtttgttct  tttgttttt
45121    gttttttaa   tggttatgta  agtgaagtag  attataaata  ggcacataca  ggatttcaaa
45181    actgtaacaa  aattaagaaa  aagctggttc  aatgagStta  gattctatga  gattaatctg
45241    aaaagggaga  gtagttatga  gaagtcttaa  aaaagtgggt  gtttgccaga  gaaacaagcg
45301    gacactggaa  gatttctgat  taagttggca  gaaatttatg  ggtgcttaaa  atgccttgct
45361    ttattcatgt  attagaagcc  cttgcctttt  tgcagtttgt  cttaagttgc  tataaactac
45421    tctctgttat  gttgaatggt  gcctgaatag  tgagaagcca  aaataattta  gttcttttca
45481    aagaagaaat  tatagactag  cttattttaa  taacccagtc  caaattataa  aaagaaaagc
45541    tttaccaScc  taatctctgg  tatagagaat  gttctctttt  tttagttgac  attgggggga
45601    Raaaagcttg  actttgaagt  tcagcaagtc  tggtttcaac  cccagccaca  actgggtgaa
45661    ctttgtcaaa  ttacttaaac  agtcttatcc  ttgggttttt  atttgtaact  ataaaatttg
45721    aStgatgtta  atgataagta  ccacagagta  ttgtactttg  aatgaaaggt  cttaattgtc
45781    aaatgaacaa  ctgaattatt  ttataggaag  agttcaactt  taataaatac  tcctgcagaa
45841    tttcatcatt  atgatgttaa  aattgatgga  agctttattt  tttaaactga  gaaactaatg
45901    cttacgtggc  ctagtttaag  aatgaatgct  ccaacttaat  acatgtttta  aaaagatgtt
45961    ttgggaYacc  aaatgtaaaa  gaattccgag  tttgtgttat  tttattgaac  aaacctattc
46021    agcacttggg  gaatcgcagg  cagcattgca  gactcgctgg  ctcctggctg  tgatctcaca
46081    cccaaactgc  aataatatcc  tgataaaacg  gacttccaca  agtaggtcaa  gaataataat
46141    gtgcagacct  aattcctcta  atttacggcg  aacatctcat  ggtcaatctc  cccgcggtca
46201    aaggactgtg  tatctcttcc  tctcagagcc  actccacagg  ctgaaaggct  gactaaaccc
46261    ctgacagata  agagattcaa  gagtagtccc  agcccccatg  gggagcaaag  acgctggctc
46321    aaaaaacgta  ctgataactc  tccctttttgc agctactcct  ttctgaacct  tccacaagca
46381    gtggcagttg  gattctgtga  ttagtctagt  cttcagagcc  agccttcttg  agttcaaatt
46441    ctgacttcac  cccttaatgt  gaagtgacat  gggcaagttg  cttaatctct  ctgtgcctca
46501    gtgtccttgt  ctgtaaaatg  ggcatcataa  taatagcgcc  tgccacattg  ggtgagtgtg
46561    agaatgaagg  aattaataca  tgtaaatcac  ttagactgtg  tctggcatag  aggacattct
46621    aaagaaaagt  tagctattat  cattatatta  ttatatgggt  ctggaattag  ttcctgaatc
46681    cttctgagat  gtgatgactt  ataaacgtag  gttgagttta  ctcatgatag  ggtcatcgca
46741    actatgcata  gctaaaatca  aattttgctt  ttcaagtttg  ttttacctgg  agccctagag
46801    ttcagggtta  tggtttтctt  tgtcactccc  cttgagggaa  gcttcttagt  cacactctcc
46861    ttctctttct  ctgcactcta  tgcactctag  aaaagctcct  tttttttttt  cttcatccag
```

95

```
46921   gcagagaggc ctactgggac ttaaatccaa ggagctgaaa tctgttttgg gatggggtgg
46981   agtcacattc tggaacctag acagagaatt tctaagttcc agaaagtgct gcttacttcg
47041   catttcctct cccccacctt tgcttttgaa actcctggca ccaatgctgc caaggctggc
47101   ggagctttcc tgagtggtgt ctgccaaatg aggagtcaag gaatatctgg aaaggcagcc
47161   tccaggtccc cgatgtcaag accatttaga actgaaagtg tcccaatatc ggggtacagg
47221   caataagcat tagttattaa tcagcctgag aagttgattc taaaatagga ggaaatgatt
47281   caattatttc ctctcaaggg attactcaat gttgttttta tgtttaaata tttatttgtc
47341   aacatcaaga attcttagta catgatKcac cagcatttt gaacaagtca tagatttggc
47401   cacaaatcaa atttcaggat gggaggagtg tctcccctt aaaatagaag agagtgagta
47461   gtctatgagg agggacctac aaagactgga aactattctt agctccgtca ctgactccaa
47521   gttcatcccc tctgtctttc agtttgggta agcatcaatt acttatctaa aatttgtaat
47581   aagaaaagtc ttcataattc atgattgtgt ttatctttat gttagtaaat ttctatgttg
47641   gKtctattct tcccccttata ttttaagaaa agaagtaaga aagtaaatga ttatttttct
47701   gaaagaaaaa attaatgtat ttattattat tcatagacct tcaactactc ttagaaagcc
47761   tttggtgact cctgggaaca agctgagtca ggaacaaaag tttgagggtt tggggttgag
47821   aataggacaa ttacttggct aatgctttag tctagttctg attttgctac tccttagttg
47881   ctttgtgatg actacttgca catgtgtctg tttgtgagtt attttgtga gcatgctgct
47941   accagcctgt gtggatgtct gtggtttcac atgatacatt ttatttcta aagtctactg
48001   aatttatgg ttttttttc aaaaatagaa actcatattc tttccttctt aatttagttt
48061   ctttcaggac attctcatga gactggtgaa atataactgc gaagtagcat gataactgtg
48121   aaataRcatg atcattgtga aagctatatg acccaaaact tccataggaa taagtgcaat
48181   gcttgcatat ctgaaacctg ctgttacatc tctttgcact taataaatat gtgttgactg
48241   attcattctc ctgggtggtg ctgagaaagt aaaaatcatg gctgataaaa attctgttta
48301   tggttttgtc taacttattt ttcagttgag atataggcta ctcttcccaa ctcagtcttg
48361   aagagtatca ccaactgcct catgtgtggt gaccttcact gtcgtatgcc agtgactcat
48421   ctggagtaat ctcaacaacg agttaccaat acttgctctt gattgataaa cagaatgggg
48481   ttttggatct tagcaattct cacaattctc atgtattcca cagcagcaaa gtttagtaag
48541   tattgccttc taagtataat ttaaattttt ataaattaaa taatttcaag aacatttacc
48601   ttgtttttaag cagtttgctt ccagttgttc cagttgaatt gtaaactgaa aaatatattg
48661   cttttgtaca atcatttttc attacaaatc tatcagaaga aattctttgt aatgaatata
48721   tcatcagttt ggctttgcat attcaaatgc agcagaattt aatgtggtgc ctgcacatgg
48781   gaagtgttgt ggatgtttat tgactgacac acagtttgtt tagatagata aggctgcttt
48841   tccattggat aactgtgggg gatttagctc caaattaatt cattaacatt tcataagaga
48901   tcttaacagt tacttattct ttcccaaggt atcacagtat atgaccggct tctaaattta
48961   atcctagaag aaaatattga gagtataaga attatgatct tttcatttga tcatttcagg
49021   attgtcttta tatctttat ttgcaggtaa acaatcatgg ggcctggaaa atgaggcttt
49081   aattgtaaga tgtcctagac aaggaaaacc tagttacacc gtggattggt attactcaca
49141   aacaaacaaa agtattccca ctcaggaaag aaatcgtgtg tttgcctcag gccaacttct
49201   gaagtttcta ccagctgMag ttgctgattc tggtatttat acctgtattg tcagaaggta
49261   ttatgcagaa ggcNcccatc ttctttcacc ctgctcccct ttcttcagtg gttgattgcc
49321   tgagctgccc ttgctttcat tccttcccta gtcctttctg gaacagttaa atttataaaa
49381   tgatttgaat aaaagtgatt tggataaact tctaggaata ctatcaggtt gaKgtctagc
49441   tcattctgag ctatttggat ttacagttgc agggattgat ttgtagctga cttagagaaa
49501   aacctagctt tcctagtgac caagataact gagagcaatt gcttactttt ggctggaatt
49561   aagaacaaac tagcacagaa aatagtaatc tggatgtttt ccatctcagK gggcctctag
49621   taggtgaaaa ggggcttcta accttcaagt taagccacag aaggctgatg agattgtgtg
49681   cctaaaaatt aattactttt gttcacaaat tgttaaatgt tttgattttg tggctgtatt
49741   tggcacacac aaaatgtcaa atgaattaaa tcaaaaagca ggatgtatta gttaagggcc
49801   taggtctgag actagacaag tgtcaaattc atgctctgcc actagttaac tgtgcgatgc
49861   cagcaagtta cttctctgcc cttcagtttc cttctctgta aaatgcatat aatgtaataa
49921   tatatttcat ctcaaagcat tattgtgaaa ataaactaag gtaatgtgtg aaaagtgctc
49981   atcatcatca ctggtacaga gtaaactctg aataaatagt aattatcttt attacactgg
50041   gattacagca ttacagcaga tgtagtgtat gaataaatgg tgaagaagtc attgttaggg
50101   ctactatggg ggctatgttt ttagctcaag tgtaacaaaa aagtatttaa actctagatt
50161   ttaatgttta tttttaaata ataaaataac cattatgtat attgatggtt tttaagaaga
50221   aaagcaatat taagtaaagg ctgaatttag attaagttat ttcacaatgc taagtgactc
50281   ttttaattgt ctgacttatt ttaacagtcc cacattcaat aggactggat atgcgaatgt
50341   caccatatat aaaaaacaat cagattgcaa tgttccagat tatttgatgt attcaacagt
50401   atctggatca gaaaaaaatt ccaaaattta ttgtcctacc attgacctct acaactggac
50461   agcacctctt gagtggttta aggtaagaag aaatttggaa ggaaatagat gaaaattaca
50521   caattaaaat agacacaagt ggccgggcac agtggctcat gcctgtaatc ctagcactt
50581   gggaggccaa ggcaggcaga tcacttgagg ccaggagttt gagaccagcc tggccaacat
50641   gaggaaaccc catctctact aaaaatataa aaatcagctg ggtgtggtgg cacacacctg
50701   taatcccagc agcttgggag gctgaggtat gaaaatcact tgaacctaga aggcagaggt
```

```
50761   tacagcgagc ccagattgca ccactcactc cagcctaggc aacaaacatt ctgtcaacag
50821   ataaataaat aaaagtgaag aattactgag aaggaaatgg aatttcttat ttcagaattg
50881   tcaggctctt caaggatcaa ggtacagggc gcacaagtca tttttggtca ttgataatgt
50941   gatgactgag gacgcaggtg attacacctg taaatttata cacaatgaaa atggagccaa
51001   ttatagtgtg acggcgacca ggtccttcac ggtcaagggt aagctactga cattaatgag
51061   atagaatact acgtgaaaga agtcgaagtg ggaacagcgg tgcccttctg gttgggtttc
51121   ttgcacttct ccctcctccc tttacttcct cctgctccat cttatcttat acattctgaa
51181   ctatgacgca aagaggtttt ctgaacacac tatcaagatt taagaaattt caggggaaa
51241   ttacattact aattcaaagc cacatctgtt cttttattctt tYtttgtgac ttaattttcc
51301   aaagataaag caatctgaat gctaacttaa cttacttttt ttgaatggca atacaactat
51361   ttggagagca aaaccagctt tttttttttt tttctagttt ggtgtcagag tttctgcaaa
51421   ttaaaaaaga gcttaatctt tagtaatact cattggattc aaagtYtaat gagaggcttt
51481   gtgatggtat actatggtgt acataaatgt tgtcgagtgg ttttaatct ttgtttgcaa
51541   tactttcaac atcatcaatg gccttgagta agtcacttca ttctaaaaat gtgtttcca
51601   agttatttta aattttataa aagcttattt aagggaaaga tttcacaatc atagcttatc
51661   aatctacaaa ggattggggt ctccttagca caagtcgatc tacagacgta gatgtaatag
51721   ccccttacg tatgacagtt ttttcagggc aaagcaatat tggagacaaa tttttggagt
51781   ttttcaatac tccaccgctg taaaataagc atcaccagac cacattccta tcagtgcctc
51841   tttctgttta atatcaaccc ttacagtggt ccttaatcac actgtcatta aataaatgag
51901   catgaaggga tggaggattg cagcagtgct ccataagcac tgcccgtctt tcagccttag
51961   tggtcacagg agtcagaatt ccgtacgggg aagatttcac tgaggatggg ccaccctagt
52021   ggagaactgc gagcaaatct gtggactcat ccatttatta ttttcatggg tctttttgaaa
52081   tcttctctgt agtcttattc ttattctgta gaagagtagt tttctaacaa ctactaggtc
52141   atgtaattag ttttatgggt tggatctgca tttgtttaag tgatatcaga gaataatgat
52201   attaaagagc atcataggta ataaaagaaa gttttattta agtgcctttc tgtttcgtgt
52261   gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgttt gtcattatgg gttattgtca
52321   gaagaacttg aaaaacattg ctatgaaata gaatagaaac atgaaaatac aagctttata
52381   ttgactagca ttcaatgctc tcctaatatt tatatttctt tttgtcttta agatgagcaa
52441   ggcttttctc tgtttccagt aatcggagcc cctgcacaaa atgaaataaa ggaagtggaa
52501   attggtaaga aaatttatca gaatgctgta aatattgcct ggaaaaatcc ttccatatga
52561   cccctgttct gaattcccct agcagggggtc aggcaattag cataaggaac cttgaggagt
52621   aagtgaggtg acatccctga aagcacctgc cccaagcatt tgctaatatt gggaacaggg
52681   acacagcaat tgcagtgttt acatttgttt attgtacttt gtaattcatg atgctttcat
52741   gtatgcatct aatttcatct tcatctctat cccagagctt gggatggaga cctgcagggt
52801   gttcattctg ggcaatggta gccagatccg gtaaaacatg tttatcttca aagtagctta
52861   tggagagatg aagagagttc tgtagaaaga tgtggaagag ggcagttgga aagaaactct
52921   aatttctagt agagggcaat cctttactt gaaatccttt gtaatgtggg gttggtgaag
52981   gcagaatcat tggccttgtt agtttcccat gcagatgaga atatagtggg agctgagctt
53041   caaacccagc tgggtgaatg aaggtaatgg aagcagggag gaggcaagag aggacataga
53101   aagaggaagg tgctagagat gagggaggga ggtcctggtg gggtgcatac taagtgttca
53161   gtaaggtttt ttttttacat taaatgggat aaaatgccag tcgcagaagt taattttatt
53221   ggtgaatgtc cttactcccc tctaggaaaa aacgcaaacc taacttgctc tgcttgtttt
53281   ggaaaaggca ctcagttctt ggctgccgtc ctgtggcagc ttaatggaac aaaaattaca
53341   gactttggtg aaccaagaat tcaacaagag gaagggcaaa atcaaaggta tttttatatt
53401   gaagagaacc atcctcttcc ccttgcacat ggtttgcacc tgcaaagtag gcattaaaag
53461   taacaggttg ctttcttagt ttcagcaatg ggctggcttg tctagacatg gtttttaagaa
53521   tagctgacgt gaaggaagag gatttattgc tgcagtacga ctgtctggcc ctgaatttgc
53581   atggcttgag aaggcacacc gtaagactaa gtaggaaaaa tccaagtaag gagtgtttct
53641   gagactttga tcacctgaac tttctctagc aagtgtaagc agaatggagt gtggttccaa
53701   gagatccatc aagacaatgg gaatggcctg tgccataaaa tgtgcttctc ttcttcRgga
53761   tgttgtttgc tgtctgatct ttgtagactg ttcctgtttg ctgggagctt ctctgctgct
53821   taaattgttc gtcctccccc actccctcct atcgttggtt tgtctagaac actcagctgc
53881   ttctttggtc atccttgttt tctaactta tgaactccct ctgtgtcact gtatgtgaaa
53941   ggaaatgcac caacaaccgW aaactgaacg tgttctttg tgctctttta taacttgcat
54001   tacatgttgt aagcatggtc cgttctatac Sttttctgg tcataatgaa cactcatttt
54061   gttagcgagg gtggtaaagt gaacaaaaag gggaagtatc aaactactgc catttcagtg
54121   agaaaatcct aggtgctact ttataataag acattgtta ggccattctt gcattgatat
54181   aaagaaatac ctgagactgg gtgatttata tgaaagagg tttaattggc tcacagttct
54241   gcaggctgta tgggaagcat ggcggcatct gcttctgggg acacctcagg agctttactc
54301   atggcagaag gcaaagcaaa ggcaggcact tcacacagta aaagcaggag cgagagagag
54361   gtgccacact gaaacagcca gatctcatga gaagtcactc actattgcaa ggacagcatc
54421   aaagagatgg tgctaaacca ttcatgatga actcacccccc atgatccaat cacctcccac
54481   caggctccac ctcgaatact ggggattacc attcagcatg agatttgggc aggaacacag
54541   acccaaacca taccacacac attatcattg ttaRactttg taaagtattt aaggtacatg
```

97

```
54601  gaacacacgg gaagtctggt agctcagccc atttctttat tgcatctgtt attcaccatg
54661  taattcaggt accacgtatt ccagggagcc tttcttggcc ctcagtttgc agtatacaca
54721  ctttccaagt actcttgtag catcctgttt gtatcatagc actggtcaca ttgccttacc
54781  taaatctgtt tgacagtctg ctcaacacga ctgcaagctc catgagggca gggacatcat
54841  ctcttccatc tttgggtcct tagtgcaata cctggcagct agccagtgct cagctaaata
54901  tttgttgact gaataaatga atgcacaacc aaattattga taccaaatgt tttttttgtg
54961  tacatttcta cttctctagc tataagtctt aattatacaa caaaatacta tttttatatt
55021  tatgtttggt aaaattcaata actttcctca tcatttggaa agtcaaattg tttattgctt
55081  ccctacagtt ttttctgaat ctagcaggat tttaatgata tcattataat ttgacacaat
55141  aaaaggacaa catgaaactg atgaatcttt attgggttaa tttcagacac tatataatct
55201  tttaaaaatg taacattctt ttttatatat aaataattgg tggcatcaca aatagccaaa
55261  gcagggtgga gagagtgatc cttcctgggt gcaggcaaga aggggatatg ttttctacag
55321  agttttcaaa acagtgataa agctgtctac aagtcattgt gcttttttatc atcactatgc
55381  ccagacaatg tgaaacatca gagatgaagt gctcttccca cagaggtgga ctgatccttc
55441  tccccactcc cttggtgtgt ctctgaatgc aatgttgtct tggaaaacag ctttccaagc
55501  atttcactcc tgagcacttg ccagtttcct cacttgttct tcacatatcc aggcaaagac
55561  atcctgtttg ctatatgaag cattgtatcc cgtataaaag gaaggaaaga gagaaatata
55621  tttttacact catcactcct caggggctgt acaatcatgt agaaattgtt taatgtgcYt
55681  gtcaaatagc caaagagtgt taaaccctga gttcccaccc atgtgtgtgg tatggttagg
55741  attcatccag atacacagag agaggcacaa caggaggaga aaggatagg gtgtgggggac
55801  agcgggcccc caatatggtg taatcgtggc aggtctctgc ctgaagtgct atgtggggtt
55861  tttcttgttt taattttgac tttaacccct gatttgtaag ttttttcataa aataaacaga
55921  atcataactc atgtagatgg ctataagtgc cgtagtgttc tgtgggtctc tggtgtctgc
55981  cagtgataag tgtggcaccc caggaaggct gtggaccca tcaaggtgct atgtgagggc
56041  catgcttggg gtggtggtgg gcccagtaga ccctgcagcc atccatccag cctgcccacY
56101  cacactgccc ttgtgtactc ctgctttgct acgttatcat tgatcaatgt ccctggttac
56161  ctatgtgttt gaattatctt cRtgttacag gtgtttaatg attttgctcc ttctagctta
56221  tttgtatttc acctgttttt cttttaaatca acatggttac actctgtttc agcaactgta
56281  taaattaaac acaaattatt actactgcta ttgagttgtc atgatgaatt ctttttattt
56341  tctgaaatta tagcatttct tgaatttaag agaacaaaaa cttgaaaggc ctatgtctcc
56401  attttattaa ctattactaa atacatattt gatgctcata ataataatac acatttattc
56461  attatttcct atgtacaagg actcatctgc ttatttaca tttaactttc tcatttattt
56521  tctcaatact tcaaagtaaa agacaaagaa agttgaataa cttgcgaaat actacagtta
56581  tggagcaagg attcaaacac agccagcatt ttcctagcta tatgtgtata cagaaagtaa
56641  tgttttgtca tcacatacct gaattactta tactttttata aaataattca cacttacgaa
56701  gacttcctct gatgtcttgg atcaacttct ttcctctact tggaagcgtc cagccaatgg
56761  catggttact tcccagcaat ccccacagga agtacacatt cctctgtgca tccagctggg
56821  gattttagag agagagtgac ctggaaagga atcctgttga aatgatttac ttacggagtt
56881  ttcattattt aacctgatga cagtaagctc tttgtcaatt ttcacttttt cccccccaatt
56941  ttgtgtcaca tcaccttgat aattcttgat tccatactgc tggtcattga gagaactgat
57001  aaactattag aggttgtgga ggaattcgtg aatatgggcc agtgattttt ctaccttaaa
57061  actgggagcc catgcatgga gacttaagac ggaaaagaac cgtatcagca aatctcattt
57121  gagattcttc tcacattcat cagtgcatat gcatgttgca tctacatttt gtgaagcaag
57181  gagtactagg aaaaatttct gggttggttg agtaatgttc accatgagag aagtatgggc
57241  tatgagtagg acctagaact tagtaatttg acttttgaat tctttatcaa agacgcttta
57301  gccacttagg tgactcttgc ttcttctcct acctcccaca ctgctaagcc cttagaggca
57361  acaagtgtct cagacccgtt gatgccttct ccttaatgcc ctgacagcac cttgtccgtt
57421  cctttctcat tattgcatga gttacaaccc tctgcatctt gcaacccaag taccacactc
57481  cattcttctc tgagtgccca tagtttcttt ccacagaaac aaactttctg ctaccatcag
57541  attaatctct tgggaaaacc actcaagtct ccagtgctcc tgtgaccttc agaataaggc
57601  aatctgcctt agttgggggtt tcaagtctgg acagaacctg ccctgacca aactttctaa
57661  tcatatcttc caccaattcc ccaaactggg ttattccatt tcctcagaaa acacacttcc
57721  tctctgcact taggcttatg gtattctctc tgcccaggtg ggcatttcct tttttgcatc
57781  tccctgaatc ctactcatct ttcagaggcc tcctcctcct ccaggaaggc tcccttttccc
57841  agtcctgcat ggagcaagct ctccttctaa actcctctgc cagcaaaaat catgctttg
57901  cctcccatct tagttaagtt gacaagtgtc tctctgcctc ttttgccttt gagaatgccc
57961  tataatgctt agcctgatgt gttgtagagg ctcaataaat gtgactgtcc tggtacttaa
58021  agcagccatg aaaattccca gtgggtatat cttatgtgaa attcttgtac tcccaaggga
58081  agccaacatc catgtatcag tttattatat cttgtttcag tagtaataat aatcttttc
58141  tttctttga atagttgatc atcatagcat ctactgcata attgcagtat gtagtgtatt
58201  tttaatgcta atcaatgtcc tggttatcat cctaaaaatg ttctggattg aggccactct
58261  gctctggaga gacatagcta aaccttacaa gactaggaat ggtaagtggc aaataccaag
58321  ttttctccc aaagaaaaag tcccataata actgttggtt acctgtctat taatctttca
58381  gtagctaggc tgctaagccc agattccatt ttgcttgcta atctgttatc agtgagttgt
```

```
58441    gtttttggat  tttcttaaag  gccattttcc  atcctgctat  gtaaatcctc  acggtcctga
58501    gatccatctc  aacagctcac  ttttcttccc  cgataggatt  gctattccta  ctgagttccc
58561    aataacagaa  tctgccccaa  gcccagaaag  gtgtaaattt  cataatgtat  cggtaagaca
58621    ttatgaagtt  aaacacagta  gcaaaattgt  tcctgttttc  cagatgtgag  caggttagaa
58681    aggtccatcg  gaatgcatgt  gtgttctacc  tgaaaccctg  atctgtaaag  tctaacccag
58741    gcacacaagc  atgtgcatga  aaggagtcag  ttctgcctgc  tagatgcaaa  cttcatgctt
58801    acatttggag  gcagatacat  gttgcatttt  tcaaacagtt  gactgaagta  ggagctcatt
58861    gtttgatctg  tggaaaattc  agatccctta  aaaaattcta  tgcccttgct  actattagct
58921    ttaaattttt  tcacttcccc  aggggagtat  ggggaaacca  aagatgaaac  caaagtctat
58981    tcagcacaga  aggcccctt   agtcatagat  tattaatcaa  aattgatgag  atggacaggt
59041    tcttgtccac  tttttttgtta ttagtctgt   gacagtaaaa  aggagaaaca  ctttgggatg
59101    aagactgtta  tttcctgata  gtgtttgtg   gcagtggttt  gacgtcaaca  tctcttgagt
59161    ctagaatatt  ttggaggatg  acatacattc  accaacagcc  ataaaacttg  gagaggaatg
59221    tcttcaaaga  gtccagtgag  aatttttaga  atcaagtaag  aagttgtact  tcttgttttc
59281    attttcagat  ggaaagctct  atgatgctta  tgttgtctac  ccacggaact  acaaatccag
59341    tacagatggg  gccagtcgtg  tagagcactt  tgttcaccag  attctgcctg  atgttcttga
59401    aaataaatgt  ggctatacct  tatgcattta  tgggagagat  atgctacctg  gagaaggtaa
59461    agctattgac  atacattagg  gacagaaatt  catgcttatt  aaaggctgtg  aactaggtgg
59521    ccttatccct  gcattggata  atgaattgca  tttactacca  caggccttaa  gaccagaact
59581    ttaaatattt  atccagaagc  agacacttat  ccttcaatcg  cccctctccc  atcattgtcc
59641    tggtgatgag  atcttcacag  taatgttggt  ggtgtgaatt  cagagttaga  agtcctttgc
59701    ttcaaggact  tagcaaacca  tcatccctac  tttattccct  tggtccccca  ccaggataac
59761    tctgccactt  cttaattctg  tccataagat  ttgaaagagg  acttaaaaat  tgatgaRttt
59821    tgttctggta  gccataggca  ctagctgaaa  taccttaaaa  gtactcagag  agtcttcatg
59881    actttcttta  tgttggtgaa  acatttgcaa  atttatattc  tcaccagaac  aaaaagaact
59941    atagcttctg  ttccttaata  ttcctaccca  attttatata  cttttgacag  ttttatcaca
60001    actcctgtgt  ttgcagtagt  tatttttccc  ttcccttcat  acatttatgt  ccagtacctg
60061    tggaccctcc  ttgtgaactc  ttctctaatt  tctcacagac  caggaagggt  ttggaatata
60121    actgggctc   agctatcaag  ggcctaagca  taataaagta  aatgttcagt  tttacatttt
60181    aaactcattt  tactaagaag  aggaattcac  atgcctatag  atgtaaaggt  atgaacaaca
60241    ggtgactttg  gtttaccctg  gaatatctgg  gaatgctaat  agcctcaata  acggctcaag
60301    agacttgtga  aagatacaat  ttaggagaat  ctaatgcatt  tcttctctca  aggctctctt
60361    cttcccccat  tttcccactg  gccacccatg  caggtagaag  taatgagtaa  tgctctcaaa
60421    cactcttttc  atattacaaa  tgtactcaag  aacctgcagg  gactcctgtt  gtttgtgagc
60481    tcaacaatgt  gccacatctg  gccacaagct  cttcacctct  tctttttcag  tcYatccacc
60541    taagttctag  ttacacttct  cctcagccaa  acccagatgc  tagctctctc  acactcaagc
60601    ttgtgctgcc  ttcaatgtgt  gacaacatac  taatctgaaa  tggttttcta  cttccYtctg
60661    tcaatgcatg  tcctctgacc  tttcctcaac  caaaataaat  gaaacccaac  aacaccaaac
60721    ttcactcttc  acttaggaaa  gttccatgact gactaagta  ctgtctatag  tctcatgcaa
60781    cttaccttgc  atttgcccca  tcaaaatcct  ctctgcacca  ccattccatc  cagtggccct
60841    gggtggctga  tcatcatgga  atccattgca  cccttgtcct  ctgctccagg  ctgggtttgg
60901    ccaatgggag  gcactggcag  gaggtcagag  ggtggtcatg  tacctgcctc  tccacgtggg
60961    gcttctggag  cagctgtgct  tctccttctg  ggagcacctc  ttctcagctg  gcctgcctca
61021    tgcatctgca  atccctcaga  cacccgtctg  tgagaatatc  ttggattgtt  ttattcacgg
61081    ttgtgctttc  tcctgttgca  gtgaggactg  ggtatttcat  gttctttctg  agtccaccca
61141    gagctaattg  gaatcaccat  ggtgtctacg  ctgtataaat  gccaaatggg  aaaggggatc
61201    tattgtccct  tgagattttt  ctagtttgga  tcaaatcagt  ttctacaaca  gagattaatg
61261    ggggtgaagt  aaagtcagga  cttactttcc  tagagacttt  ggacaatttt  gaggggcaaa
61321    gcttcctgag  tttgtgattt  ttaatatgtg  agttaaaatt  ttaagccatt  gaaaaacata
61381    taagcaaata  ttctaagaag  ataatatgta  ttctaagaag  aatgaaccag  aattctaagt
61441    ctctaatgca  acaatagact  aaatatcttt  acattcccat  ttacttggaa  ggactctgtt
61501    ttagttaaga  ctaagctacc  tgggaaataa  aaagaggaca  ttggggctga  tattctcaca
61561    aatgaaagga  acacaaagaa  caaaacgggt  tctctatcca  cacatacttc  cacttctctg
61621    agtaagtgac  ttgatgtcag  agaatctcac  accagattat  aacaataaga  cttttaaatg
61681    ttcaggatgt  ttatgtttaa  agcattagac  tgataagaaa  tctgatctat  ttcttgtatg
61741    actagatgta  gtcactgcag  tggaaaccaa  catacgaaag  agcaggcggc  acattttcat
61801    cctgacccct  cagatcactc  acaataagga  gtttgcctac  gagcaggagg  ttgccctgca
61861    ctgtgccctc  atccagaacg  acgccaaggt  gatacttatt  gagatggagg  ctctgagcga
61921    gctggacatg  ctgcaggctg  aggcgcttca  ggactccctc  cagcatctta  tgaaagtaca
61981    ggggaccatc  aagtggaggg  aggaccacat  tgccaataaa  aggtccctga  attctaaatt
62041    ctggaagcac  gtgaggtacc  aaatgcctgt  gccaagcaaa  attcccagaa  aggcctctag
62101    tttgactccc  ttggctgccc  agaagcaata  gtgcctgctg  tgatgtgcaa  aggcatctga
62161    gtttgaagct  ttcctgactt  ctcctagctg  gcttatgccc  ctgcactgaa  gtgtgaggag
62221    caggaatatt  aaagggattc  aggcctcagg  tttcatctgg  taactttcct  tccatttccc
```

```
62281    tgcttctggg  tggatgcaaa  atggccagaa  attttctcc   tcaatcctcc  accatcctcc
62341    agccaccatt  ttcttccctt  ttcccttttg  cttgcctgtc  ttcactgtgg  tgtgggatgt
62401    tttcactact  cttttttcct  ctctctctgt  cttgctttgt  ttctcctcat  tcttgtcaat
62461    aatctgaata  gcaaactaat  cattcgagag  tagattttca  cgtcacttga  agaacattct
62521    gattccctca  ggcagaatgt  cagttgtaaa  tcttgttccc  aacatgccag  attttcttta
62581    ttttctatat  atatatataa  tattttatat  ataatattat  atataatatt  ttatatataa
62641    ttttatatat  aaaatattat  atataatatt  atatataata  ttttctatat  aaaatgtgta
62701    tataattata  tataattata  taaaatataa  tatagaatat  ctaataatgt  ataatatata
62761    acatataaaa  ataatattat  ttaatatata  atattttata  tataatattt  ttatatataa
62821    tataatatat  attttatata  taattattaa  ttatataatt  aatatataat  atatattta
62881    tacataatta  ttaattatat  ataattaata  tataatatat  cttatacata  attatcaatt
62941    atatataatt  aatatataat  atatatttta  tacataatta  ttaattatat  ataattaata
63001    tataatatat  cttatacata  atatatataa  atatattata  tataatatat  attatatata
63061    atattatata  taatatatat  tatatatata  aaatttatat  ataatattat  atataatatt
63121    atatattta   tatacaatat  gatatataat  ataatttata  tattatatat  atttatatat
63181    aattattata  taaattatat  aaatataaat  tatatatta   tatataatta  ttatataaat
63241    cattatataa  ttattataat  tataatatat  aatataatat  aatattatat  ataatatata
63301    gtattctata  taaataatat  aacatatatt  ttatatagaa  tattatatat  aatataatat
63361    atattttata  tagaatatta  tatataatat  aatatatatt  ttatatagaa  tattttatat
63421    atataatatt  atatatatgt  atttgagaca  gagtctccta  tgaacatact  tgtggactta
63481    tgattttatt  ttgacactta  agagtggtca  tagggtaggt  gcatatttaa  ctttcagaga
63541    aacttataca  ttgtttcgca  aactaattgt  actactttac  attcccacca  gttgtgtatg
63601    agagctctgc  tgcttcattg  cctcaccaat  aattggtgtt  gtcaagtttc  tttatatttt
63661    aacaattttc  acaggtgtta  gatgaagatc  accatggtct  taacttgaag  ttccccgact
63721    catgatgatg  agccctttc   atgtttcat   tggccattcg  tatatatttt  ttagtgaagg
63781    caaaatattt  tgccttattt  aaattgtatt  tcttttata  ttgttgatgt  gtaggaatta
63841    tttatgtatt  gtggatgcaa  gttctttgcc  agatagatgt  aaagtgagaa  tatttccagg
63901    gttatctatt  cattttaag   tgatgtcttt  gtttagtaga  cattttaat   tttggtgatg
63961    tccaattaat  caatgtgtcc  caaatggcta  ttttttttg   tgtcctagct  aagaattttt
64021    ttttttgcc   tgtccaagag  catgatttat  ttatttattt  tttgcctgtt  caagggtctt
64081    ctctgtattt  ttatgaaata  tttacggttc  tagcttttat  gtttaggtcc  atggtctatt
64141    gtgaattgat  acttgtatat  agaacaaggc  aaggacaaag  ttctttttc   tgctgctgct
64201    tcttctgtag  aggtacccag  ctgttccagc  aatgtttgtg  aaagactgtt  ttttccccac
64261    tgaattgaca  tggtgctttg  ttaaaaaaat  caattgacta  gatatgtgtg  gatctattag
64321    tctactttct  attttgtatg  tttgtatgtt  tctgtgaata  tcacattgcc  ataattacta
64381    ttgccttatt  ataaatcctc  aagtcagtat  aagtctctca  actttgttca  tctttgtcaa
64441    gattgcccag  acaggccggg  caggtggctc  acccctgtaa  tttcaacact  ttgggaggct
64501    gaaatgggtg  gatcacctga  ggtcgagagt  tcaaaaccag  cctgaccaac  acggagaaac
64561    cctgtctcta  ctaaaaatac  aaaagtagcc  aggcgtgata  gtgcatgcct  gtaatcccag
64621    ctactcgggc  ggctgaggca  ggagaatcgc  ttgaacccgg  gaggcggagg  ttgcagtgag
64681    ccgagatcgc  gtcattgcac  tcactccagc  ctgggcaaga  agagcaaaac  tccgtctctg
64741    tccggacaat  cctagattat  ttgcattttc  atatgaattt  tagaaacagc  ttcttaattt
64801    ctttgaaaaa  ttttcctggg  atttggattg  aattgtatta  aatatataga  ttctgtaata
64861    tatagagtct  aaaaatatac  agtgtggaga  atagatgtct  taagtcttcc  aatccgtaaa
64921    cgtggtatat  ccctttatta  cttagatgtt  tcatctctgt  tagcaatatt  ttgtagtttt
64981    cacaggagat  atctcatata  ttatccaatg  aatttatact  ttggtatttt  gatttttaaa
65041    tgctatggta  aatggcattt  ttaacaaact  tcacttttca  agtttccatc  tctaatatat
65101    agaattgcta  ttgatttta   tattccatga  tcttgataaa  cttaatattt  ccagtagttt
65161    ttttttttgt  agatttcaaa  agttggtcta  catataatta  tgctctctgt  taaaaccagt
65221    agttttaatt  ctttcttctc  aatctttatg  ctttctaatt  ttattttat   ttcttcactt
65281    gttgcaaagc  agttgtccca  ggtgggagcc  aaaatcctta  tctcattact  aatcttagga
65341    aaaagtaaaa  tgtttcacta  ttatgaggtc  acctgtaggt  ttcttataga  tgttctttat
65401    cagattaaga  aagatcattt  ctactctcgg  ttttctgatc  gttattaatt  cagaaagggt
65461    gttgaatttt  gccagatgct  tttatttgca  tttattgaga  taattacata  tttttattta
65521    ttatgtggtg  aattatctaa  attgggtaaa  aagaattaaa  atcattgtcc  aattgaacat
65581    tttgcctgtc  ggctatgggt  ttccctttc   ctttggttaa  ataacagttc  tgccacaaaa
65641    taaaaatcta  gaaacacaca  ttcccctttg  gcctccttaa  aaaaaattaa  aacttcaaca
65701    attgcacccc  tctctattac  ttccctatag  tgaactccca  tccttctggg  aaagcaaaaa
65761    gcctggggcg  catttgtgta  cttgagacgc  ggagcccctc  tcaccttcgg  catcccccgg
65821    ggacccctct  tccaccctca  gccctcccct  cctcttcctc  tgaagagccg  ggcggccgct
65881    gcggccagcg  gtcctccctc  ctggcatcct  ctgctgtgaa  ccgcggcctg  caggggtgctt
65941    gcgagcggcc  cgggcgcgcc  cccttccggc  cccgcaggtc  acccggctac  agcccgggtt
66001    ttcccgggac  ccgcgccgcg  cggctgggca  ggaagcgcca  gacgcctggg  gcccacgccg
66061    tccgcgggaa  aagggcaagg  cgctgggttt  tccagcagca  acctttggac  cccgcgatcc
```

100

```
66121   agtagctccg gtaactccac gcggggcgtc tgggtggagg agccggtcct ggagcacggc
66181   tgcgaggagc acccggaccg gagggtcccc agaccgggac ctccgagtca gggaggattc
66241   tacgccaggg agcgccccag actgagagcg ccccagaccg ctacactccc ggacccggag
66301   cttcgcccga ccgcgggcag tgcccacctg cagcctccac cggccggggt tagcagccag
66361   gagctgccag acgcctgaca ttcttctttc tgttcctact ttttttcctt cttcttcttt
66421   tttttttttt ttgtagccct ctctgggtgc cttatctctt taatcacacc tctctttcac
66481   tttccacggt agtcaggagg cggagatcgc tgcttctcac ctactttctg aacttggcct
66541   ccgcagtcgc gacctggcgt gaaggaggag ctgccgcccc cgccccagcc tcggggacgc
66601   ctctctgaag gtaagggtgg gctccgctgc cacccgcatc ccctccccac ccccagagg
66661   aagggaagac caccatccgg ggagggattc tcaaaacaag atcaattaag agagagagag
66721   agatagtgtg tgtgtgaata cgtgtgaata attgatgaga cagcatcctg tctttctctg
66781   aaagtatagt tttagaaatt aagacatccc caattatttt ctggagaggt gagaacgact
66841   gaaataaatg ataacaatag ttgtcatttt ttaagtgctt actgtgtgcc tgacactgta
66901   gaaaatttt agatatctta ttttcactta atcctcctaa gggccacatg aaataggaac
66961   tattattatg ctttaataaa agggaggagg aattccgtca tttacccaag agtatccagc
67021   tactaagtag tagatcccag atccacacct acagtttttc tgattccaga gcctttgtct
67081   ttttctgcc cagccttggg aagttaaaat gcccagcacc ttgcatcgca gttatcagac
67141   aaattaatga ggtgtcaatt gcagccagat acagagatcc aagattattg tattcatttc
67201   actgacttcc cttgcttcat tctcaagacc ctttgcaaat actgattttg tgtcatacac
67261   agtgctgcta ttgggagcac acctctcaac ctcagggcct tggatgggct aggaggggta
67321   aatttgggga attgggtttg tgcaggttct ggggacagtg ttttccctct ccctgcacaa
67381   atatacatcc aaaggctgag aactaaggct aaggcctagg aatgtggatt gaaaaggcag
67441   ggtctaaatc ccagcacttg ggaggctgag gtgggcggat catgaggtca ggagatcgag
67501   accatcctgg ctaacatggt gaaaccccgt ctctactaaa aatacagaaa attagccagg
67561   catggtggca tgcatctgta atcccagcta ctcgggaggc tgaggcagga gaatcccttg
67621   aacctgggag gcagaggttg cagtgagcag agattgtgcc actgcactcc agcctgggcg
67681   acagagtgag actctatctc aaaaaaaaaa aagaaaaaaa agaaaaggca gggtccctgc
67741   cttcagggac ctcaccagag agtgggagct agtccaggag tgatggaaac acctgcagag
67801   tgatggagac gggcttggac ataggcatgt gccatttggg gctgttgcag ttggagagct
67861   tccttgagaa ggcaaatact gagccaagtt gaaaggactc tagggacctg gcatctggga
67921   taagtgggga gggtggccca ggtagcacag gcataggcag atttgggggg aagcataagg
67981   ctaacaacag acacttgtgt gacaaggact atgaggtggc attgattttc ttggccttac
68041   agatgagaag ccagattcca cacaaattat tttgtttgtt agagtttaag cgtgtgtgtg
68101   tgtgtgtatg ttatgcccca agatcatttt ctcctctttg catttggttg ttttatcatc
68161   tcacaggctg agtaaaacga ctatgcatgg aacactgtag aagaaaaatt atattaaaaa
68221   aaaactggtg ttatgatttt catccaggca ttttgtctca ggagttgctc gattgtatgt
68281   gttgctaatg agtgatttat attgaggcat ctacattgat tcatgtgtgt ccttgtgaat
68341   agctgtaatc tgaggctggc catttcactt cctcaaatct caggttcctc attagaaaga
68401   gctttgaaat ggctctttcc ctgtacatta tgtgtttctg tgaattactt ttagagaaag
68461   gtcagtttct cagggtggga cccataaaca agggcaactg ggcagagcca ggggttcagg
68521   tgactgtttg gaccctctgg aagggctttt tactgtgctg ccttgggaga cagagtggtc
68581   agcttggaac tgaccttaag cctaggtcac ccaatttttg aaggaaactt ttttttttca
68641   tttagagaaa acttcatgcc tttgggtgat taggtaagtt tgtgactatt cgttatgtgt
68701   gtgtacacac acgtacgtac atatgtacat gcacaaatat ttaatatgtc ttagacacac
68761   caaatatatg catttaagat gtattaaatg tactttcctg agaatacaaa taagtgccct
68821   acgtcggcgg ctctcttcct gactgtatgt tagagtcacc tgcggagttc tgaggactca
68881   agatgcccag tctgtcctcc agaccactgg gtggagcccg ggcatcatta ttttttttcca
68941   aacttgggtg agagctggtg gcctagactt cacattagtt gctaaaacat actttccaag
69001   atatcatgcg ttaaaaagca tttcaaaagt agattgtgca gatgtgccaa tggatcagat
69061   gatcacctag attcaaaatt atgtatattt tatggatgta attcattcaa taattcagca
69121   aataacaaag aaacaagcag acataaagtg aaaatatgac agttagtgat aagtgctatg
69181   aatagaaaaa ctaaacatgg gagggtagag gaacattggt gtgtgtgagc atgtgtgtgt
69241   gtgtgcgtgt gtgtgcatgt gcatgtttgc atttggaaaa gcatgctgtg ctagacttgc
69301   agtagagtgc tcagggagga cctgtacatt ataagctagt gtcagcaaac cgtaggggaa
69361   atacaggatt atttaattaa caatagagaa gtaattagta ggtatttgag gaaagtaagt
69421   ttagttactc cctatcttca tgtcaatcaa gaaaattaat tgttggtgga tgaaaaaagt
69481   taaatataaa aacaaaacaa aacaatactg agccataaaa atgaagaaac tatgtaagca
69541   tctatttgac caccaacagg aaaggatatt ctacacttaa aagcaatgaa aacatctcaa
69601   agcaaaagaa tgatcggttt gactacatca atataaaatt tgtcatatat tagcataaca
69661   aaagagagaa agtaaacaac aagttagaca aacatctgag aactggacaa aagacataaa
69721   gagatggttc ataaacatga aaataaaatg gctgataata agtgattaat atttgcctca
69781   tttatcccca aagaaatgca aattcagaaa acaatgagat accattttc cattttggct
69841   gtctaattac attttaaaag taaagctgct cagtgtgctt ctgggtgtgg tcgatagatg
69901   ctgtcttgtg ctattgagtg agagaatatg aatattgcaa cacttctgga atgcagtttg
```

101

```
69961    tcagcatgca tctgtcaggt ttcaatgtag tcttcaactc aatcactgta ctgccatcga
70021    gtaacatcca ggaaataata agaaatgtag acaaataata atagctaaca tttaagcggt
70081    attaccattc accagacact tgcaatagta actcatttaa cttcacatc agttccagga
70141    gatggtatta ttatcccctt ctatagataa gaaaattgag gcacagaaca ggcaagtcag
70201    ttgcccacag tcggttagct ggttgaaggt agaatggaga ttcaaatcca cccaggctgc
70261    cacctgagcc catcctgcaa aaaaactgag ccatcctggc ttacctgtga agatggtcat
70321    gtgattatct cataatacat ctggtaaatg ataatatatc cacatgaagt aatattaggt
70381    ggcaattaga atggttttaa caaagagttt tgtgacacag gaaaaatgct cattatgtaa
70441    tattaggtgg aaatgcagga tatgcaatct tttaaacatc ttttattttt tgagaaggag
70501    tcttgctctg tcacctaggc tggagtgcag tgatgagatc tcggctcact gcaagctccg
70561    cctcccggat tcacaccatt ctcctgcctc atcctcctta gtagctggga ctacaggtgc
70621    ctgccaacgt gcctggctaa ttttttgtat ttttagtaga gacgggtttt caccatgtta
70681    gtcaggatgg tctcagtctc ctgacctcat gatccgccca tctcggcctc ccaaagtgct
70741    gggattacag gcatgagcca ctgtgcccgg ccaaacatct taattatgtg acaacattta
70801    taaaaagtac aagaaaacat gttaaaatat tatacatgtt gagattattt tcttctttca
70861    gtcttcattt tgtaaatctc ctacattgag tgtaaattgc tttggaaatc agaaaaactt
70921    taaaaagaaa gaaacttaag agttacggaa aagcaaaaat attgatgtta aacataacaa
70981    ttaaaaaccc ttaaaattct tacatgtgag acatttaata tttatattct tctatgccta
71041    ttttattaat acttgctaga gtatctgtta catttatcag aatatctctt tttctggaaa
71101    tgaattcatt atttaatata accttaatct gttaaattca cacataaaac ttaagtacat
71161    taaaaatttt acatacacac atccacattt gcagaatggt gtgcaaatac cagctaattt
71221    ttgtatactg agatcttaat atgtttcata ttgaaactca gagttaactt gtcagtttgc
71281    tgagccatct cctgataact ttgtctcttg gtagatattg aatatttttt ttaaagtggt
71341    gcttttttata aggttacatc agtcttaatc agagtgacat aaaggaaata acattttta
71401    ttctatccaa gataaaagaa taggcagagt taataataaa atgtgtgtgg aggtagtggg
71461    ggaggtttca aaccatagac ccagtctaat ttgattatgc acagttggta ctttccattc
71521    atacttgttt ctttggcctt tacattattg attattatct cttggaaatt ccaccaaata
71581    tttggtgatt tttccataag gatgtatgtg aatggtttgg cagcttagcc attaagcttt
71641    atcctggatt caggcattga caactctgtt tttaggatgg aagatgagta gttgttggtg
71701    agagggcat tattctgagc tgagaagctg ccatgcactc aggaagtctc cataagcctt
71761    agatggttta atcactcaat aattccagaa acaggccttc ttatcatctc cattttggta
71821    atgaccaaat aggctctaag acttgctcaa gttttgcaca gcttgtcagg actcaaactg
71881    gagtggcagc ccaggtttgt gtgtttctga aaccagactc tgccatcagg ctgtgctccc
71941    tgcactcagg atggtaaaca agggatatgg tgttgtttat gcagtattta gactctgaat
72001    ggtctccact gtgggggatc cttcctccac tgggctcaaa gttgtaatca gaactccgaa
72061    tgctaggttg cagtaggatc aaagttcaat gcagatctta gttccactgg gacacagtca
72121    atgtaaagga ggggtctgga gtgtaggaga ctgtcacttc atctctccat ttccctaaag
72181    taaactcata tgggcagcca cagacacaYg cgaacacatg catacacacc tatgcctcag
72241    tgacacatgt tagtgagtgc tggccatgtg ctaggcatcg tgctggattt tagtgcatca
72301    gtgaagagca aaatagactt agtgcctaga attatgaaag aattatgaaa cctacgttct
72361    gaaggaggaa gatggagttt taaaaattca accctaaatg aagatctcat tacaattcag
72421    gaatgttttg atgtgaaaac acacacacac actctgttac atgggcagaa ctgccataca
72481    tagctcagca tgtcacagat tgtttattca tagaatggta atatttgtag cagtactatt
72541    gttctgaact ttggccttgg aagcaggatt tagtagacca ggcctgaaat tcttcagact
72601    ctgttgtatc agagaaataa attattaggg atttcttcca gaagaaaaac ttcattgcct
72661    tgaaatttta ttaattttat atggctgaaa gttctgagaa ttggttacta tattaaggat
72721    tataaaaatc ttctaggttg tttttttaaa aatctgtgtg ccagaagatt tttaaacYtt
72781    cataagatag gcacactttt gtttgaaagc tttggctgaa tctgtttat tctgtttcc
72841    agagaagcca tttgaagcag aatccaaacc atgaattgta gagaattacc cttgaccctt
72901    tgggtgctta tatctgtaag cactgcaggt aagtgattat acatactctc aaacatattt
72961    catgagtaat tgaggaagaa tgtcaaagtt ttttttttta tgtggtggct actgaagatg
73021    ctgaatttat tcagaaaaga aaagattctg cagtctgagt tcatgtagtg agagcagagt
73081    gaaaaacttg catcttgaaa taaattacat ttgtctgagg tgaaaattta tttcatgtta
73141    tttgaaactg aggaggtggc agccatcttt attattgggt tttaggaagt ttgaaacaca
73201    ctttctcctt cctgagaggg aggcaaagat tcatctcatt atttaggtaa taatgagaaa
73261    agtaaacgac aaatatatta gacagcttaa aaaaattaga gtgagacaga atatacaaaa
73321    aagagactgg ctagaccaag gagctcaagt aggcttcata gtagatgtca aatgccaatt
73381    ttaatttatt agcagtgatt gcctagagtt tttttttttt atttgagaag gattctgaga
73441    cttcattctt tagcgttcca tgattgaata gcaaggcatt ccttaggctt tggggagaa
73501    agtgggagca caaatattaa tttgccttga tctattatct aggagaggga aacacatggg
73561    ctcatttccc tactttgata attactgctg aagagatagc tgaattaagc ttttggtact
73621    ttcaaataac acaggtccaa tctcagggta tatttactta cgtgaatcca tttattattg
73681    agatccctgc aatggaaccc tacccagatg ataaaagaat gaggtagatc ttcattttt
73741    tatatgtaat gatctgtcaa acaagtgagt ggaaaagcaa ggtacaggaa aatgtgtcat
```

```
73801    gagttccatg tcctgtggac atacactaat agccttactc ctcattataa tgagagcatg
73861    tttcaggcac agaaacagcc atgcagccag ttgtgcaaag ttgccaagct cccagctgcc
73921    tgccctgaag agcacttttc atgaatgttt cactctaggg agacatgaca atgttctaca
73981    tgtaggattc actgaaccta gttgagttta aatgacaaaa tgctatgatt gtccaaactc
74041    aagacagttt caaacacaga ctttattcca aagggcattg gttgttggca ctgctgggtt
74101    cttatactgt gtttatggaa caataggagt gcctattttg gcacttggaa ctccacaaac
74161    cttgaattcc acttatttat cccatcaagg accatttcca acaagcatta tggtaatcaa
74221    caaagcctct aatggtaaac catgattaga tatccaaagc ttctagctgc ctgaggaggg
74281    gggtggtgtg gaStttcaat tatcaacctc ccacctctgg agagtgtgta attcacatta
74341    agtttcttca aatggtaata cttccagaga gcacatgtcc atccaaacac agtcatatgt
74401    gtttatagaa gcttttttaag acaagcgaaa acctgaaaac ctagaaataa gccaatatcc
74461    aacagacatt tgtttaataa accaaatccc caacatgcat cctttttataa aatggataaa
74521    ttgcgaaatg gtaacacgat tagatactac attacaatga gaaaagaaca catcactgct
74581    acagcatgaa tgcatttata acaatcatgt caaaataacc cagacataaa ggagcacata
74641    ttgtatagcc tatgcatttt attatattaa ctctaaaaca ggaaaataaa tgaactgtta
74701    atgactcacc agggtttttgc atgtcaattg aagctgacac ccacatatac acatgtatca
74761    caccttctat gttgaagtat tagtataaaa tagagattgt attgcaacca gaaaacatga
74821    agttggtgct ccttaactct tgtgtttttgt ggtagattcg acatccatgt tttcaaagaa
74881    tgtgtattag tccattctca tgctgctaat aaaggcatac ttgagactgg gtaatttata
74941    aaggaaagag actttattga ttcacagttc tgtatggctg gggaggcctc aggaagctta
75001    caatcatagt ggaaggggaa gaaaacatat ccttcttcac atggtggcat caaggagaag
75061    gctgagcaaa aaggggaaaa gcccccttata aaaacatcag atcttatgag aactcactca
75121    ctatcaggag cacagcatga gggtaatcac ccccatgatt aaattacctt ccaccaggtc
75181    cctcctatga cacatggggc ttatgggaac tacaattcaa gatgatttgg gttggggtat
75241    agccaaatca tatcattcca ccctggcct ctcccaaatc tcatgctctc acatttcaaa
75301    acacaatcat gcccttctaa cagtcccca aagtcttcac tcattccagc attaactcta
75361    aaatccaagt ccaaagtctc atctgagaca aggaaggtcc cttcctatga gtcagtaaat
75421    tcaagagcaa tttagttact tcctagatac aatgggagta caggcattga gtaaatacac
75481    cccttccaaa tgggagaaat tgactaaaac aaaggggcta ctggccccat gcaagttcta
75541    aatttaatag ggcagtcatt aaaccttaaa gttccaaaat agtctccttc aactccatgt
75601    ctcatatcca ggtcacgctg ctgcaagagg taggatccca tggccttggg cagctatgct
75661    cctgtggctt tgcagggtac atctttcttc ctggtttctt tcatgggctg gtgttgagtt
75721    tctgggcttt ttcaagtgca tggtgcaagc tgtcagtgaa tctatcatcc tggggcctgg
75781    aggatggtgg ctttcttctc ttaactccac tagctgtact ccagtgggga ctctgtatgg
75841    gagccccaac accacatttc ccttccacac taccctagca aagattctac atgagggttc
75901    cacccctgca gcaaacttttt gcctggacat ccagcatttc catacatcct ctgaaaccta
75961    ggcgtaggtt cccaaacctc aattcttgac ttttgtgcac ctgcagcctc aacaccatgt
76021    ggaagctgcc aaggtttggg gcttgcaccc tctgaagcca cagcctgacc tggtatcttg
76081    gcctgtttta gcaatggctg gaggggtagg atgcagggca ccaaatccct aggctgcaca
76141    cagaaaggga ccctgggccc atcccaggaa atcatttttt cctcttaggc ctctgggcct
76201    gtgatgggag ggtctctcat gaagatctct gacataccct agaggcattt ttccccttgt
76261    cttggcaatt aacatttgag tcctcatttc ttatgcaaat ttctgcaacc agctttaatt
76321    ccatgccccc cattcccacg aacatggatt tttctctcct accacattct caagctgtga
76381    attttccaaa ctcttatgct ctgtcacctc ttgaatgatt tgctctttag aaatttcttc
76441    ttccagataa tcgtctctca agttcaaagt tccatagatc tctaggacaa gggcaaaatg
76501    caactaatct ctttgctaaa cataacaaga ttcactttg ctccagctcc ccaaaagttc
76561    ctcatctcca tctgagacca cctcagcctg gactttattg tccataacat tatcagcatt
76621    ttggttaaag ccatttaaca agtctttagg aagttccata cttttccata ttttcttatc
76681    ttcttcaaac tgttccaact gttccaaact gccctccaaa ctgttccaac ctctgcctgt
76741    atgcagttcc aaagtcactt ccacattttg tgtatcttta cagtagcacc ccacttctgg
76801    taccaattta ctttattagt ccattctcac actgctaata aaggcatacc tgagaccagg
76861    taatatataa aggagaaaga tttaatggac tcagttctgc agggctgggg aggcctcagg
76921    aaacttacaa aaggggaaa agccccttac aaaaccatca gatctcatga gaactcactc
76981    actatcatga gaacagcatg agggtgactg ccccccatgat taaattcaag atgaggtttg
77041    gatgggacca cagccaaacc atatcagaat acttctatca catcacttcc tctttatgtt
77101    gtggttaaca attttgagtt aaaaccagac tgctgctttg agaaacaaac cctgactata
77161    gagaaggcat tctgtccatg tcctaaggga aatgtttcgt tcagtttaaa tactctcatt
77221    ctcaaaatgt ggtaaaagtg cacccgagag tcactggagg gtagattatg agaaagaccc
77281    attaatgctg aacttgaagt accacaagtg ccgaaagagg ggaacgagca cctaYcctgg
77341    gtagatgatg gtcagatata tttccaaagc cagagaatct gggtcaattg agaaagaggg
77401    aaagtgacag atatgcaaat ttggattgat attatattaa tgaattcttg attatccaca
77461    tggtgaccta cagtaataat aaatgtacac acactgtaag ttacataggc acacttttaa
77521    ccatcccccac acttaccagg gaatgttaga taggcagggt gaatgcaagg tgagtgagat
77581    ttctgatggt gcagaggaag aacgccaagg ttgctagatt tcttgttttc ttttctactg
```

103

```
77641   agactcctca gcagggtata ttttgaggta ggtatggcag caccgtgact tctgtggctt
77701   tccctatgaa gaagttgtaa tgtgtaatat tttcttctaa gaggtgacac aattatcagt
77761   cttctagcta tggttggtta agttctgata ccttcctgtt ggggagaaac tgtcagaagt
77821   tgtgttttca tctatgttct gaaagtgcct tatcctttct agtgttggaa agattgacgc
77881   aaaaataaag aacccacagg aaattgacac cgagatttct caggactttc ttgctaacct
77941   tgcttcttca cctaatgcat taggagacca aaaaaagtta ccttgtcatt ttggtttttg
78001   tttttattta ttttactttta ctaatctttYt gaagaatctt gtacttcacg tccccacatt
78061   actgtggttg aaggggaacc tttctatctg aaacattgct cgtgttcact tgcacatgag
78121   attgaaacaa ccaccaaaag ctggtacaaa agcagtggat cacaggaaca tgtggagctg
78181   aacccaagga gttcctcgag aattgctttg catgattgtg ttttggagtt ttggccagtt
78241   gagttgaatg acacaggatc ttactttttc caaatgaagt gagtaaccct ttctttttcaa
78301   aatgtatttc acagccctct tgtcctttgt tcagcaactc aaatatgcag taatctgaaa
78361   ggtcagataa tcaScacatc cttttctttc ctactcttcc tatgacatga aatacattct
78421   Wtgttatgga acaRaataag tttatctctc tctgcttatt ttcttagtga ccttaatgaa
78481   aggagattac tttgatagga tttccttttta ggggatcctt accaatgaca agtgacgatt
78541   gcatgaaagt cagtttttaaa attcagctgc tcctgttgtc tacatgcccc cagtcaaatg
78601   aaatcagcag ctcggacttc acgataatat taatgtgttg cgaatggggc agtgtcagct
78661   tcctgggtgg catccctggt ggggagcagg gatggggttg atgcagaaac accctagtcc
78721   agtctgcagg ggctcactgg gagccaggtg gcagtgctag tggaggaagt gtgtgtctgg
78781   cagccaatcg caacccacag ctggcggcag tgagagaggg caacacggaa gggctttgca
78841   ttttactcat gaggtggccc tagattgtgc cagaatactc gaaattgggt tgcataactt
78901   ccatgaaatg agcacagcca actgaaagca tcacccacac ttaccgaaca tgtttctttt
78961   tagatgtacc tacctacata tgtaggtagt tttatagaga cttgagactg acctgtaaga
79021   aatagctcag ggccagctga aggtgatggt gtgtagattt catatgtagc agatttcttt
79081   tagaagcact tctagagtga gtcctgatca ttatcaacca tgctgttgag agcagagtga
79141   gctaccactg tgaaaccttg gtagcacttc tgtRgttttg caccaaatca ggtcattttt
79201   gtttgttttt atacatagga actcccaggc atcagaaaaa aattaaataa ctcttttgat
79261   aaaaatcagt acatctttag atgacctttt aagagtgata attacagaaa tgggtcacaa
79321   ctgggatttt tcataacaag ctaaggaaga atgaatactc acaaatgggg taaaggaaaa
79381   agaacaacaa acaccgNccc ccccaccaca aacttccaac tcctgattga atgaaggcac
79441   aatttaaagg aataaagtat tctggacagg cctgagagag cacgttgaga ggaacagcca
79501   ggcatctatg atcatcactt ggagataaca gaagcaaatg gcattggcca tctttctgat
79561   aYgggtagta gagaatacag cctggcttag gaagcaacaa tggaagcaag agattgaggc
79621   aggatgtcag agagaggtgt ggagagacag ggtagcagat aggagatgct. aagaaaaatg
79681   aggtgtccca gagtggataY tggagactga tcactatgag agtaaagtgW tgaggagaga
79741   agcagctttg acaatggcct tgaaaataaa tgggattcta catggaagca ggttgtttta
79801   gacacccagg agcatagaag atacagacat attcaaatat agcataatta gtaagcagta
79861   tttactgcag atgtgtgtgt atacacatat gtgttctaac ttaRgtgtgt aaccctttgY
79921   agtttgatgt gggatctRtc agtgaacaga ctttacacct gaaaatttcc cRttagaaca
79981   gtaattcttt aatttttagg agtacattgc taataatcaa atgccactga ggactgataa
80041   tttaatcctg cagtgaaaat cactttatcc caccaaattt taagaacgaa atcaactaag
80101   aaggaagggg atacaaaagg gataRttttc aatgatagag atgaagaaat gaaggtagag
80161   ggaaatgtac acacagaaat ctaataaacc aatcctaggc tgacaggcag caagttcagc
80221   ttggtgagca gcctcagaag tgggggctgg tgtttatccc ctggtgtctg tggacaaaaa
80281   tgagtggttt tcataaccag gatgggctgg gcctgcagca ggaaagtgtg gtcacagctt
80341   tgggtcagtt ggcctcagtg ttaacccYag ccttggagct cctggatggc aggtgcagtg
80401   tgttgttgat tgtttcagtg ttcccagagc agtgtcctgc acatggagtg tcatcagtcc
80461   tctagcattt gttgagtgtt aggaattgag gattacagag gtagagaaca atgtagtctg
80521   aaaaaatggg gatattcatg taccctattc cttttaagca acaagtcata gagatccaga
80581   aatctctatc tctccaattc attgaggtca aagttggggga gacattacga aaacacactt
80641   gagaagaagg tgctgccaat gggatcctga atccatttca agagccaggt agctctatca
80701   tgtcccatgg ccatggaccc actgcctata actgtgacaa cagcctgcca cattctccag
80761   tcacggccaa tcttttataa attcatgccc aaggtgaagg gccagattc attggattgt
80821   gaaagtgtag aattgtgcag agcacttctg gcctctgggg ccaaatagat gtgcatttgt
80881   atcctggccc catcttgtag caagtgtgtg gcctgagtga gttgtttaac ctgtctgaga
80941   tttagcattc tcatctgagc ctaaggaaaa tagtcatctc ccttgcatgt acatagtcta
81001   atacctggaa ctcagtaagt gttagtcttt tcttccttct atcctattgt agacctaaag
81061   gggaactgga agagcctaag gtgggcacaa atgtgtcagg catagaggat aacttaaaaa
81121   cacgcatctg ctgtatttag caaataaatc tatacatcac aggtgcagtg gaagcaaaca
81181   tgaaatttga aacattttct ctttaaaatc ttcatttaaa aaatactgta ttgtttttcca
81241   attccagatc accttttttt gttgttgttg ttgttgtttt tttgagatgt aatttcgctc
81301   ttgttaccca ggctggagtg caatggcgcg atctccactc actgcaacct ccgcctcctg
81361   cctcagcttc ccgagtagct gggattaccg tacccagcta attttttgtag ttttagtaga
81421   gacggagttt caccacattg gccaggctgg tctcaaactc ctgacctcag gtgatccacc
```

104

```
81481    tgccttggcc  tcccaaagtg  ctgggattac  aggtgtgagc  cactgcacct  ggcctagtga
81541    tagtcctgat  cactaatttc  tgatcaggag  taacaacaac  tataaaatct  attcacactc
81601    acttttggtc  agtaaggcat  ttaaagagct  aattcaatta  cactgagggc  agcaactcca
81661    gactttggtt  tattcaacag  taaaagaagg  ggattggaag  aagttgcctt  gagaaacaaa
81721    acctaatata  tatatacaca  catttagaaa  acaacatcag  gagtttatag  accttttgaa
81781    accaaagcca  ccctcccatc  ctgggccaag  actctctgaa  cttatccatc  agtgaagatc
81841    actgacattg  taaagatttg  cttttctgac  atccataata  tattgtctgc  agtaggatat
81901    gtgcgtgtgt  gcgtgtgcat  gtgtgtgtct  ctgtgtgtag  actgcttagg  ttttcattta
81961    ttaccccctg  ctcccatgac  agtttttaaa  gttcttgctc  ttctgaatca  agcggaggca
82021    gagcagggtt  tgcacatcaa  tgaatcaatc  tctttaataa  aaataaagga  aaaatattgt
82081    aactggttac  taaagggaag  atgggtgata  tttgcaaagt  ttctgtagca  ttataaaata
82141    cctttacact  tttattccat  agaaattata  ctcagaaatg  gaaattaaat  gtcatcagaa
82201    gaaataaaca  cagctgtttc  actgaaagac  aagtaactag  taaaattgtg  gaagttaaaa
82261    aattttttca  gataacctgt  gaaaacagtt  actatcaaac  actggtcaac  agcacatcat
82321    tgtataaggt  aatgcttta   taatatttta  actttaagac  ttgatggaaa  agataaaatt
82381    cccaaatgag  gggctgagag  ctaccattct  ggtctcataa  caaataagct  gcctgatccc
82441    acatggctca  cttgccctgt  ttgagccaag  tatcctcact  ggcttgttgt  caagagctga
82501    tacatagtat  taaaatatgt  tccacagtag  tttccctgga  tttggtgttc  aataaatgat
82561    cacagtttct  tctcccacac  tcttccttca  ttttccttct  ctttcgataa  gttagttttt
82621    aggaatgcta  gtttaaggat  ttccattttc  tcctgtcaca  tgtacttaac  atgtaaaact
82681    caattgtata  tgtttttctg  aatatattag  ttcattttat  aaaattggaa  atggtaatgc
82741    ttttataata  ttttaacttt  aaaactgaaa  aatgaagaaa  gcaaaaatca  ctcataaatc
82801    catcacagac  ttttttttttc  tgtgaagcat  ctgctttta   aaagtggtat  attttatttt
82861    tttatgtttt  ttcctctaag  aattacgttg  tgactatccc  ttcacatctt  taaatgatct
82921    tctaaacatg  ggctttcagt  ttttatgtgc  ctggtttcat  ctaactaacc  cccttgtttc
82981    gtctaactaa  tcccattgtt  attggacatt  tgtgaattgc  ctattttca   ctattattta
83041    aaggctaaaa  tgaataccat  ttccacaaat  ccttgggtgt  ttatatgatt  tgtccttcaa
83101    ccaaatccct  aaatgtagaa  ttactggatc  agaggctgtg  agcatgtttt  tatgtctttt
83161    tcctcacact  atcaaacttg  tctccaggga  gctctgccaa  ttttaccccca  cacttgaaaa
83221    aagaatccca  attttttgtc  acctccacaa  cacagggaaa  taatggtccc  agaatgttta
83281    aaatttgata  tgaaaaattt  ctaaaactcc  tagtggtttt  ttcacttttc  ttagtaataa
83341    attaaacttt  tgtatatatt  tattgcatct  tactgcaatc  aggcaatttt  gtgggttgcc
83401    tgtttatttt  tttctgtttt  cttcagttgg  tatacctttt  acctgttgaa  ttttaaaagc
83461    tcttatacac  catggaatac  tatRcagcca  taaaaaagca  tgagttcatg  tcttttgtag
83521    ggacatggat  gaagctggaa  accatcattc  tcagaaaact  aacacaggaa  tagaaaacca
83581    aacactgcat  gttctcactc  ataagtggga  ggtgaacaat  gagaacacat  ggacacaggg
83641    aggggaacat  cacacaccgg  ggcctgttgt  gtggggtggg  agggtagggg  atgtatagca
83701    ttaggagaaa  tacctaatgt  agatgacaaa  ttgatgggtt  cagcaaacca  ccatggcacg
83761    tgtataccta  tgtaacaaac  ctgcacattc  tccacatata  ccccagaact  taaagtataa
83821    taaaaaaaaa  agctctttta  gaccacgcta  taggaattat  atatacacat  acatagcgtg
83881    tgtgagtgta  gtgtgtgtat  gcataaaaaa  catacgcact  cccatccaca  cacatacgca
83941    acacaaattg  tgtagttcag  gggtccccaa  tccccatgcc  actcccccctc  actaacatta
84001    ccacctcctg  tgagatccgc  agctgcatta  gactctcaca  ggagtaagaa  ccctattgtg
84061    aactgtgcat  gcgaggaatc  tatgttgcat  gctccttatg  agaatctaat  gcctgatgat
84121    ctgtcactgt  ctcccatcac  ccccagatag  gaccacctat  ttgcaggaaa  acaagctcag
84181    ggctccactg  attctacatt  atggtgagtt  gtataattat  ctcattatat  attacagtgt
84241    aataacaata  gaaataaagt  gcacaataaa  tgtaatgcgc  ctgaatcatc  cccaaaccaa
84301    ccattccccg  ctcaacccct  gcccaccctg  gtccatggaa  aagttgtcat  tcatgaaatc
84361    agtctctgtg  tcaaaacggt  tggggacagc  tgatgtagtt  tatatggaat  aacaaatctg
84421    aataatatga  ctctgtcgtg  gaaactttaa  agacctataa  gagagtggtc  tgaacagaac
84481    ttgggaaatc  tgaagcgagt  gatgataata  tccttgaaat  ttttgtttgt  gccttatgct
84541    tgaacactgt  ttacttgtga  tttccaaaga  cagctcttat  ctgaaatgtt  ttctaaacac
84601    tcgaatcagt  cctgagttga  ctttattcta  tttcccatca  atgtacaaat  gtgttacagt
84661    atatgtaaat  tagtgaggct  tggtgtgctg  agttttaaga  aattccgttc  gggtgcagtg
84721    gctcacacct  gtaatcccag  cacttgtga   ggctgaggtg  ggcggattat  taggtcaggt
84781    gatcgagacc  atgctggcta  atatgatgaa  accccgtctc  tactaaaaat  acaaaaaatt
84841    agctgggtgt  ggtggcacat  gcctgtagtc  ccaactgctc  gggaggctga  ggaaggagaa
84901    tcacttgaac  ccgggaggca  gaggttgcag  cgagccgaga  ctgggccact  gcactccagc
84961    ctaggggaga  gagcgatact  ccgtctcaaa  aaaaaaaaaa  tgaaattcca  aagtgcagct
85021    atttaaatta  tgagcattgg  tgcaccaagg  atatttgatt  aaacaattga  ccacagggct
85081    aagtattcta  agcctgttga  ctgatcatga  cttcacccga  gagcaaggag  gctaattcag
85141    tagataattg  attggaaagg  aaaatgaaga  gtctctgatc  ttttctatag  cttgaaaaaa
85201    ttagtttggc  tttataaaat  tttgacttat  acataaataa  tttactgacg  tgaggtacac
85261    atgtgggtgc  ttcacgggca  cattctatgc  acaatgggtc  tagccccagt  ggctgccagc
```

105

```
85321   tctgagggag  aagtctgtgg  gcagtaggtt  ctccttgctt  ggtgtttgca  gaaacagtgt
85381   ccaatggtca  aggatggcca  ttcagccccc  agacatccca  tgtggtccct  gcatggctct
85441   gggacctacc  tctgggtgtc  agatgtgctg  gattgaagaa  agtggcttta  ttacaattag
85501   tattaattcc  tgtgacatgt  atggaaatgt  gtgtgWgaga  gagagagatg  tttgagtgta
85561   cctatggatt  tgaatattgc  tgatgcattt  ctaatcagct  cgctggcttc  tgatggaggg
85621   tgtcaaaggt  gtgggtggtc  tgctagctac  tcattctctg  agtaattctc  cacccatccc
85681   actatccacc  ccttggatgc  atagtccaga  atacacaaaa  ttggactgaa  catgtatgaa
85741   atagagttgt  gactcaagca  agggtttttg  aagtagtgtt  ctctctctct  ctctgtctct
85801   cgctgtttct  ctctctctgc  atgtgtgtRa  aagagagaga  tgaaagaata  ttgatgaaag
85861   cagctacagt  gctcatattg  ccccatgctc  tacttcccag  catatattaa  taatgataat
85921   catgccaaca  gcaatgggat  cgcatttaat  tgtttcagta  aaagtagtac  caagtaactg
85981   ggacatatgt  ccaggaattg  ctactttcca  ttctttttc  cttttctct  aaagatcttt
86041   aactcaaaaa  tttggatcac  tgtaatatca  atttggcttt  tacttaaagt  attttaactt
86101   gttttattaa  aaacagaact  gtaaaaagct  actactggag  aacaataaaa  acccaacgat
86161   aaagaagaac  gccgagtttg  aagatcaggg  gtattactcc  tgcgtgcatt  tccttcatca
86221   taatggaaaa  ctatttaata  tcaccaaaac  cttcaatata  acaatagtgg  aaggtaaggg
86281   aaatcttaga  attgggaaga  aacagacgta  ttttagtaaa  atggaatttt  ttcattcttc
86341   aaaactgtgt  aggaattgag  agtccactta  gttgggggttg  tgaccacacc  actagttcaa
86401   atatgatggt  caaaatcttc  attcYggtgc  tcctctactt  ggctaaataa  acctttcttt
86461   ttggcaggaa  atgaaaatct  gtatgaataa  acctcacatg  agtcagctct  ctcttatata
86521   tgaaatttga  gtaaattgtg  gaatctttg  ttgttgaatt  tacattttca  ctttacaagt
86581   atttagtgaa  gatatctata  ctcaataaac  tgtacttagc  acaaggaggc  agggaagcaa
86641   ggaagtatga  gatgaagtga  gcactgtact  attctcaagg  gatttgattc  taatcaggag
86701   aggcaataag  ctacataaag  gtgtatgaaa  cagcatcatc  tgtgaggctg  caagagtgca
86761   gggggaggtt  ggaggaagga  cagatgcagg  aggctccatt  tgacctgggt  cttggagggg
86821   gagcctcatg  aaacagtgac  ctgctggggg  aagggatcctc  cactgggaga  gtgtaaggaa
86881   cactgatgtg  ctggtgggag  gggagggatg  atggaccag  gctgtgtgtg  atggaggaac
86941   agaatcctgg  gacattaaag  gtatttaatt  aaacctcacc  cagtgaagat  attcatcata
87001   aggcacagcc  aatgattgat  catccagtcc  cttcttggga  actatcaagg  aaatgaccaa
87061   atcacacagc  agtgtagtcc  agattttaat  atctgattgt  tgaagtcttt  ttttttccatt
87121   gagtccaaat  atcccttctc  gtagcttcca  accactgatt  ttgatttgcc  ttcaagaatt
87181   tcacagatta  aatctggctc  cttcttaaaa  tgacagctcc  tttcgtgatt  tgaagacagg
87241   ttgcttatcc  agtctggatt  ttttttcccca  tcatcctgag  aaactggaaa  ctgtaacact
87301   tgtgtgggtg  gcttgttcaa  caaaatctgt  gggctcacat  ttgctacagc  ctttgcttta
87361   gcaacatctt  tcttcttcac  tacaaccata  tagctatata  cacccacatg  ttgccacagt
87421   aacaaattat  cccaagttta  gtgtcttcgg  taaaatcaag  gttttgccag  aattgcttta
87481   cttctggagg  ctgtagggga  aaatgcgttt  ctttatttt  tttatttatt  ttatatttttt
87541   gagatggagt  cttgctctgt  tgcctaggct  gcagtgcagt  ggcgcagcct  tggggctcac
87601   tgcaagctcc  gcctcctggg  ttcacaccat  tctcctgcct  cagcctcccg  agtagctggg
87661   actacaggca  cccgccacca  ctcccggcta  atttttatat  ttttttagt  agagacgggg
87721   tttcaccgtg  ttagccagga  tggtctcgat  ctcctgacct  cgtgatcccc  ctgccttggt
87781   ctcccaaagt  actgggatta  caggcataag  ccactgcacc  aggctggaaa  atgtgtttct
87841   ttgcctttgc  tttctccagc  ttctagagac  cacttgcatt  tcttggcttg  tgccacttca
87901   tttcagagcc  aaacacattt  gcttccctca  cctcatctcc  tttttgactc  tgacctttct
87961   gtctccctct  gatgaggacc  ttgagacagc  attgtgccca  cctagataat  ccagggtgat
88021   cccctatct  caagatcctt  aatttaatca  cacctagagg  ttcctttttgc  catgtaacat
88081   atYgtattca  caggaactca  ggagtaggat  gtgggaaact  tcagggtcta  ttattctgcc
88141   tgccacagcc  cagaacctgg  gaaacccata  actgaatttt  aaactcaaaa  acttaaattc
88201   tgtgtataat  catctaccct  aaactctcca  actttacttt  cctacatcta  ttttttaatct
88261   catataaatc  tctgttactt  tgatactttc  agaacctaca  ccctctttcc  caccatcttt
88321   gctttctgcc  catcagcagc  tctttccttc  cttctttgat  gaagctgtag  gaatatccat
88381   tattaaaagg  gtgcttgtta  tgggccaggg  atagtgccaa  gtactggagt  tgaatcatgt
88441   catataatcc  tcacaacagc  tctgtgagac  tgtggccttg  ggtgagttaa  tttccttagt
88501   tcctctctgc  ctcagattcc  tcatcatgaa  atgggacagt  aatggttcat  ccctcgtagg
88561   gttggctgac  aatgactgag  tgagtgtgtg  tacagctcta  ggatggctct  agggacagta
88621   agagttctgc  tgtaatgacc  atctcccact  gctaggttcc  caaacctcac  ctgcagcaca
88681   agctattctt  ctcccataat  tgtccatctt  atttttcttg  tcttcttgga  gaatgtcaac
88741   tgagtctctc  aggactccac  tcaaatgtcR  ctcctccttc  atgcccccat  ggcctcttgt
88801   ttgtatttct  gattgctgtc  agaattgcaa  tacatgtgtg  tctctgtcaa  taacttgaaa
88861   ggtttctgga  aaggtttgca  ctgacttctg  ttcctcatct  ttgcatcttg  gatacctaca
88921   attggcactt  ggcacataaa  ggtgctcatt  acatgcctgt  tgtgtgcatc  actgagacag
88981   aaacaactcc  tttgatgcca  catacaacaa  gaagtctaga  cacttcacca  tcctggctat
89041   ttttctttgg  atttggcaaa  gtttgccacg  agtctacaga  aattagagct  tggagttggg
89101   gaagacagcc  aggaggtgtt  ctgacctgtg  ttgaggacag  agaggttctg  agctcgtgcc
```

106

```
89161   ttctccatgt tgtgtcttta ttacagtccc cttgccttca cagatggaga tctctcagct
89221   gctacatagt gactgctgta gtggtggaaa ggtgtgatac ctctcctccc catcctaagg
89281   gtcatggcaa tgttcccata acaaaagaca gattaacaag gggcaagcaa gttttatgtg
89341   acatgggagc cttcagaaat gaagaccctg gggccgggtg tggtggctca tgctggtaat
89401   cctagcactt tggggaggccg aggcggatgg atcacctgag gccaggagtt cgggacaggc
89461   ctggccaaca tggcaaaacc ccatctctac taaaaacaga aaaattagct gggcttggtg
89521   gtgggcacct gtaatcccag ctacttggga gactgaggca gggagaattg cttgaaccca
89581   ggagatggag gttgcagtga gccgaggttg tgccattgca ctccagccca ggtgacagag
89641   caagactcgg tctcaaaaaa aaaaaaaaaa aaaaaggaa agaaatgaat accctaagac
89701   tcagggaaac ccgttttttt cttttaaat gcttaggttc tagaagcatg gacagccatg
89761   tagaaatgtg actggacaaa agggtatgac ctaatgctaa tggactgagt ggggaaaccc
89821   agcaagacct gtccagattc ttcttggcct ctctgttata tttcctttc ctgggtagag
89881   ggcagagggt cttaagactt actgtcagac caaagaaggc caaagaattt atagccaact
89941   cctagacaga aaggcaggag aaggttagag tcatgcttct aggttctatg gcttactttg
90001   gggaaaagga ttctagtttt tatgatctgc ctttgggaag agggagtctg gtttctgtga
90061   cttgctttgg ggagaatgga ggcatgaagc aggagggtca gagagatctt gcttttgag
90121   gctgctttga ggccttccaa tgtcgttgag tttaaagtac tcagcctgcc aaaacaccat
90181   actttgggac tgttttgga gcccaggagt tccattgcct tcagcttgtg ctcgtgaatc
90241   acagtcacct gaattcctac tgggcatctg ccactttcct ggtcttatgc tctggctgat
90301   cgggtggcgt ctcctcctac tgctgacaac ctctatgcca ttttgactct gatagcctcc
90361   taggtaggct ggccatgtca agggcatgca ctaagatgtt gtcttccttc catgacctac
90421   atgggactgc acccttcact ttacaccagt tgtgcaaact cagggctgat ctggatttcc
90481   aatgatataa tttgagtgtt agtagacagc atcctttgct gaagatgtac ttcacactgc
90541   cccttctcaa gtctcccatg taagtttctg ggggaaaaaa tggatccaga gaaaggtagt
90601   cagtagcagt tttcccccctc aYaccgaaat tgcctcctgc ccttgtctt tgaattcaag
90661   ccggtgttgc ttcttactgt cagacctgag agttagttct ttccctggat acatctcagc
90721   tcctcctgcc tgtttgccaa cattccaggc ttgcccagag actccaaatc ccaggataac
90781   ccagacctga ctctgaactc cggagtccac tgtgctggtt tcattctcct gttcttgctc
90841   acatatctgc ccactgcatt ctcaactgct cactctggca ctgcactttc tgagagggct
90901   caggaggtg tgggagctgc aggWgcactc ctgagtctca ctgtactcca ctcaaccagg
90961   cattctgctt gctgagcacc tgcaagtccc agaagggatg ttctggacat tgcttgggaa
91021   tgcctccaaa gcagtggggc tactagcagg gaacagtcct ggtcatgagg aaagttacca
91081   atgaggatat cagggctaac aataacttga ctatgcacca gttcaaatct atgaccgcag
91141   ctgctcttag ggtttgcagt tgccagctgc cttaagcaag agcctgttgg gtgatgactt
91201   acttcccaag catcaccaga tggggactga gggggcagaa tttatgtgca gtacttccag
91261   tcggggtcag aagccactca ggagcagaag tagaacttgg cctacaggca tgaacttaac
91321   agcagggcca gggccaaggt gaggttcctg gggcacaata ttgaagggac actgctctca
91381   gcatcctgca agggaacggt gggcacatag agtgagggat ttgctttct caacctgtct
91441   tggccctgca acctgtaaat ccatcactgg ccagctgagg aggacagatg gagagaatcg
91501   agacagggca gaaaggctct tcccagtgat tttgcatggt ttgactcttg cacgtttgat
91561   catctgagag tcaggctgca gtggatgggc atcttcctct gccttttagt tgaattgagc
91621   agaagctgca gatMtgagtg agagaaaacc agatacaagg tgctttcaac atctgtgacc
91681   tttagttaat gcaaatagga ctcaagtgaa cgtaaaccag taactattgg aatctttgtt
91741   acatgaaatg agcatattat catagatgct tgtcataaag aagctcaagg caacactaat
91801   acatttgtct tttctttgt cacttctaga tcgcagtaat atagttccgg ttcttcttgg
91861   accaaagctt aaccatgttg cagtggaatt aggtatattt caatatacat atattctgca
91921   tttataagta ggggacataa tagagccttc caagcgtaaa tatgattcat tattgaatga
91981   cactggatac acatttcata tttactgaat aacattacaa aactaagaRa agagtgccat
92041   tttttgaatg ttgccacttt taacctgact taaaagtggc aacattttaa agttcaagag
92101   acagtggaac tgcctgtggt ctgtagacta aggttttgt tgtgctgtgt gtggcttgtc
92161   tttaactctc tctatggtca tttctaaata tgaaaattta ggccgggcat ggtggctcac
92221   gcctgtaatc ccagcacttt gggaggccga ggcaggtgaa tcacttgagg tcaggagttt
92281   gagaccaacc tgaccagcac ggtgaaaccc catctctact aaaaatacaa aaattagcca
92341   ggcatggtgg catgcacctg tagtcccagc tactcaggag actgaggcag gagaatctct
92401   tgaacccagg aggcaaggtt gtaatgagct gaaatcacac cactgcactc cagcctgggc
92461   aacggagcaa gactctatct caaaaataat aataataata acaataataa atttaaaaat
92521   tccgtgagag cactctggag agatgagtca aactaatttt gtggtgtgaa tgtttgatta
92581   attgattcac tttgaatcaa ttgactgaat tgattgtata tttggagttt gtctagataa
92641   gtcttatgaa tcaacatgtc caattgagta atttctttca aaagagaaag gaaagaaaga
92701   aaggaagaaa ttcacatatg ctaaatatgc taagccaggt atcatgctag gtcattgaaa
92761   tatattgttg aattaaatcc tccaagcaac actgaggcat gtattctcat tttatgaatt
92821   aaataattga agtgcaaaaa ggctaagtaa cttgtctaaa gttaaatcac taaaaagtgg
92881   caagtctctg tttgggact ttgcttccaa agtccccacc ttataaccac agttgggggtc
92941   acagtggttt atatgaccac atgggcccag tgcagctggg tcctatgcca tttgtacatg
```

107

```
93001   aggtttacct  gctttatatt  tttatttatt  tctgtctttg  tctttcttat  ttttttcttc
93061   tactcagtgg  ggcttaactt  gtatcttttg  ttctagtttg  ttgacctgaa  gagttagata
93121   attaattta   attatcctct  ggtagtatat  accttcaaac  tccaagtttc  cctgtaaaca
93181   ctactttagc  tgtttctcat  attttcattc  tattcagttt  aaaccattac  ttaaaatttc
93241   ccttatgatt  tctttttcca  acaaaaaaaa  ttggtatttc  ttaatttcca  aatatttggt
93301   gaactttcta  gttcttgtct  cttattcatg  tccagtttag  attcattata  gtccaagatt
93361   atacacagtg  tcatttcaac  cttttggaat  ttgtagtgac  ttacttaagg  cccacagtat
93421   ggtatatctt  gctgaatgtt  ccaagtgcat  ttgaaacaaa  ggtatattct  ataggttttt
93481   gttttattat  tctttaaatg  tctatttggt  taagttcatt  aatggtgttg  ttcaaaaaat
93541   ttatatcctg  actaatcttt  tatctgcttg  ttccatgtgt  tgctgaaaga  gatgtgaata
93601   tctccagcta  ttatggtcaa  tgtgtctcta  tctcctttag  ttctgtcaat  ttttgtctta
93661   catagtttgg  agctatctta  tatgcataca  cggatttatg  ctcgttgaat  taacactttt
93721   atcattatga  aatatctctc  tttatctatc  ataacacttc  ctctctcaat  gtgtagctct
93781   gtcatatgaa  ataatcacac  aagtttttctt  ttgcttggaa  ttggcatatt  tgtctttttt
93841   taaacatccc  tttactacta  acattttttt  ttgttctata  cataagtgtc  tctcataaac
93901   tgcatgtatt  tgggtctttt  ttgaaatatg  gtcagtatgg  tgtttagtcc  atctgtatat
93961   ttaaaattac  tgttacagtt  gctatatcat  atatgcatga  tatacaaata  ttcaatttat
94021   tcttttaaa   ataaattaca  ctttgttgaa  aatgaccata  tttccatcca  cttgtttatc
94081   ttttcctcta  ttttctttca  aacagttaca  gtttttaatc  tgctaaattc  agcatctgta
94141   ggtctgcttc  tattagctgt  ctttttctatt  gattatggtg  cacattttct  gtgtctttgt
94201   atgtctctta  acttttatt   gttgttgatc  tttatgtatc  aaacaattaa  agtgaataca
94261   gtttgtgtta  ttttttctta  caagtaggta  gcctgagagt  tgagtatctc  agttgactca
94321   gaattgagtt  ggggtggctg  ggatgtaggt  atgttcactc  caatctgtct  tcatctctga
94381   tagctggtga  gggaggggt   gctgcagtca  tgtgtttgat  gcaggcctct  tctctaatga
94441   gacttgctct  cctgattacc  atgagactgc  aagtgatatc  atacagcctt  tacatccttt
94501   tacgctgtcc  gagcgcttag  cagtaattct  gcaagcctgt  gcagctgccg  agagcttcac
94561   tggatttct   tgtttctatg  agactctttc  tgttaattac  aaaccaaagc  aaattcttcc
94621   atggaatgaa  atggccacct  atcttagctc  accctagaaa  ggatccttct  taagtacgct
94681   tgtagttcaa  aaatttctct  ttattttctc  tgtactttga  tgtctttaat  gtttgcttaa
94741   tttaattttt  ttgtgattta  tccatttggt  ttatctagaa  gtctcagtga  gagtattgtc
94801   ctgatgctac  gaaaagaaag  tacactgatc  actctaagtt  ttaatagttg  attaaaaaat
94861   aaatgtttaa  aaagaagcca  aatcatgttt  tgatgatgac  ttgtaattgc  tacctatttc
94921   ttctgcttta  tatatttgt   tttatttaaa  gcaaggcata  tttattctct  ttctgggatg
94981   gagcaccacg  ttttgcttta  ggttaacata  cattgattat  attgttttga  aactttttaag
95041   aaacctgagt  ttcaaaaatc  agaactttag  ttgccaactt  ttaccaggaa  aaaacgtaag
95101   gctcaactgc  tctgctttgc  tgaatgaaga  ggatgtaatt  tattggatgt  tYggggaaga
95161   aaatggatcg  gatcctaata  tacatgaaga  gaaagaaatg  agaattatgt  atgtatgtgt
95221   aatatatatg  tcatgatata  taccatataa  caatcatata  tattatatat  aacatcatat
95281   taaaaacact  gggtggctga  tgtgtatttt  gggctagttt  aaattataat  tataatggac
95341   atattttRta  attatagtgg  ccataccta   ggtggacgca  ttatgcagaa  atgtattgcc
95401   aagatctaag  ctggaaaaga  aggtccatga  gaaggccatt  gagggtattt  agcatcacct
95461   tataaagtag  cacttggtca  gtcaggatcc  gatggggatg  gggatttac   ttagcttctt
95521   agaaaagggc  agttgcctag  gccaagcttt  ttttttatttt  gaaaaatttg  ttaagttgtt
95581   tatttttcta  atgggccaaa  aaggctccaa  atggcatgca  tattgatagt  taaataagcc
95641   tatattcttc  cattaaggaa  tggtcaaact  tataaaaaga  atcagtctaa  acatatagtc
95701   acctaagctt  ttctgaatga  ctttgggtca  tactgtaagt  attcattatt  atactgatac
95761   actactcagg  gtctagtgag  tgctgagcct  aagttgtcta  ccacctttgc  tgcacttgtt
95821   gagttttttc  ttctgctttt  tacacatttt  ttggtacctt  ctttgttacc  atgcttgcag
95881   ctctacattc  ttgctgacac  ttgtctacaa  gctctttaca  ccacaaagtc  tggcttcttg
95941   cccgggagca  ggcatggtgg  tcatgcccac  tacatgttga  attcatttga  aatttgtgga
96001   tctcctgcca  tcatatctca  tcctcagcca  ctggcctcca  tttgagtttt  gtatttctcc
96061   ctttggataa  tctctttcta  attttttagag  gggtagaaaa  ccacagtgaa  tttcttagaa
96121   atgcagctta  agctacaaaa  gcagRgaggt  cttttcttttat  agttgggtgc  acacacctttt
96181   gatatcatct  gttttgtcct  ctcaatggcc  tcctgtgttt  tacttctgga  tttcattggt
96241   tttagatggt  tgttatttct  ggatgggtga  cttaggaaaa  gaattacatt  ccaaaatgta
96301   cttgaggcag  ttggaggcta  cagtggggct  ctgtgtcccg  ggtcctagat  atccctgttg
96361   ctaggaatgt  gggcacatgg  ggcccatgtc  tggctgtgtt  gactggtgta  gctatcagag
96421   cctgtcattg  gtgggggaag  tggatgtcat  aagtgtttac  ctggcatcaa  ccacaggaca
96481   acttgaatcc  tttaaggaag  agaaaaagaa  cgtgctgcca  gtggctttct  caagtcattt
96541   tctgacactc  atggaacaac  ttaactggca  cttctcatgt  gcctcggggt  gcttgagaac
96601   agtgttatgt  tttttaatgc  tgaacacatt  aaaccagttt  gttggattcc  atggacaaat
96661   gtaatacgaa  tctgcatcta  tgctaaacaa  cttacagcct  tgtcataatt  agtgggaacc
96721   actctgtgga  aagaacccca  agttgttagg  aagagagaag  tagagaaaaa  taattgagag
96781   agaatatttt  ctYatatttt  caacttagca  tcatgatctt  ccagccattt  aaacttcagg
```

```
96841    taattgacag  caccattccc  aaatggcctc  aggagggctg  tttctgtcag  ttctgatgct
96901    ccacggagga  ggggacgttt  attgtagtcc  atgctttaag  agtactggcc  agaaagggaa
96961    gaaataatgc  aggcaacatc  acagctgcct  tcaagcattt  taaacatgtg  aattMccctc
97021    tcaagggtaa  cgaacagagc  ctactgctaa  attattttgc  cctYttacag  gactccagaa
97081    ggcaaatggc  atgcttcaaa  agtattgaga  attgaaaata  ttggtgaaag  caatctaaat
97141    gttttatata  attgcactgt  ggccagcacg  ggaggcacag  acaccaaaag  cttcatcttg
97201    gtgagaaaag  gtgagaaaga  tttatttttg  gaagttttga  aacttagctc  attgctactg
97261    agagacattt  caaagatggc  cactttctca  ttttccacac  cagaacccag  ctcctgagag
97321    gggaaatcag  tcacctcatg  tcacaggcat  tcagtggggc  ccctctgtct  catgatgcca
97381    agcagagaat  gaaaacagaa  ctcattcacc  tgctctcaca  gtcatattcc  atcccatgtg
97441    gaaatcgggg  ggcctgaggc  ttggagatac  agatgctgag  agtgaggctc  atggggaggc
97501    catcaggaca  acccctgcac  ccagcactgc  acccagagca  gccagagggg  agtttgcaga
97561    tctgcgtggg  gtgaggccag  tgcaaacttg  gtgtgactgc  tgctccagcc  tgcctcagca
97621    gggtaccagg  ctcagggtcc  ttgtgtgtgc  tcattcattc  attgcctcct  gaacttggca
97681    agcaggccat  tgatagccct  gatctgtggg  Mtcttaaagt  ttttcagtgc  tcacctgatg
97741    gagggcccca  gggtcgccat  ccctttcctc  caccaaagta  gcacagccca  gggccttcag
97801    cccagtagct  gctcagaaat  ggctgtgact  ggagccagaa  agctgtgatt  tccaKttgtc
97861    actactcatt  cactgttcat  tcccaggctc  caccatcaac  tccctggata  tgtagagaaa
97921    ggtcaagatc  tgcaagagag  agagaaaggg  agagagacag  agactaacac  acccagagac
97981    agatggagac  ggagagagac  agacacagag  agaggagaaa  gagagataca  tgagatggag
98041    ggacagagac  aaacagacaa  ctctagagag  atgaagacag  aaagagaaag  agaaagagag
98101    acagagacac  agagagacag  agaaatgaga  gagacaggga  gagacacaga  gacagaggaa
98161    cagagaaaga  gggagataga  gaggcagaga  catgaaggga  tgcaaataca  gagagatgag
98221    agtgaaacag  acagagagac  agagtaaggg  agagatacac  acagacacac  acaaacatgc
98281    agagacagcc  acacacacag  actgggcctt  ggctttgtga  gaaacggaag  gagtgggaca
98341    ggaagagcac  aaaggctgga  aaccccctcc  agctgggcta  tcctggtgcc  caggggcagc
98401    tcccccagca  cacatgtttt  agagccgcct  ctctaccata  tgcactcata  acaacagctc
98461    acatcccata  catgttaatt  acatgacagg  atctgggtga  agttcctgcc  atgatctcat
98521    ccagtcttaa  cgacaagctt  tatagaggtg  aagaaacttg  cccagggtca  cacggcaggt
98581    gaaggtattg  aaactcagac  ccgattaaat  ctggaatctg  tgttcttacc  ttctgcaggg
98641    cttgtcgcct  gtggtcttcc  acccctgtgt  gattgcctgt  gctgtggtcc  actcacttgt
98701    gagggagYgg  ggctcaaatc  tttgttggac  catcacctgg  ccaggctttc  ctggggcctg
98761    ctcaccaccg  gatcacctct  cctttcttag  aacagggcag  gactgtccta  ttcctgccct
98821    agacgcttct  ggccagtcaa  ttctcctgct  gctgctctca  caaatgtagg  tttagaaatt
98881    ggggctcaca  acccaccatc  accccgtgcc  aatgtgtgcg  tgcgtgtttt  tgcaaatgtg
98941    tgacacagtt
```

[0251] Three alternatively spliced transcript variants encoding distinct isoforms are depicted hereafter. cDNA sequences 1-3 show the human cDNA structure for transcript variants 1-3 of IL1RL1, respectively. cDNA sequence 1 encodes the longest isoform (1), which is depicted hereafter as amino acid sequence 1. cDNA Sequence 2 differs in the 5' and 3' UTR, compared to variant 1. The resulting isoform (2), depicted hereafter as amino acid sequence 2, has a distinct and shorter C-terminus, as compared to isoform 1. cDNA sequence 3 differs in the 5' and 3' UTR, and contains an additional internal segment, compared to variant 1. The resulting isoform (3), depicted hereafter as amino acid sequence 3, has a distinct and shorter C-terminus, as compared to isoform 1.

IL1RL1 cDNA Sequence 1 (SEQ ID NO: 2)

NM_016232 Homo sapiens interleukin 1 receptor-like 1 (IL1RL1), transcript variant 1, mRNA

```
  1  aaagagaggc  tggctgttgt  atttagtaaa  gctataaagc  tgtaagagaa  attggctttc
 61  tgagttgtga  aactgtgggc  agaaagttga  ggaagaaaga  actcaagtac  aacccaatga
121  ggttgagata  taggctactc  ttcccaactc  agtcttgaag  agtatcacca  actgcctcat
181  gtgtggtgac  cttcactgtc  gtatgccagt  gactcatctg  gagtaatctc  aacaacgagt
```

109

```
 241 taccaatact tgctcttgat tgataaacag aatggggttt tggatcttag caattctcac
 301 aattctcatg tattccacag cagcaaagtt tagtaaacaa tcatggggcc tggaaaatga
 361 ggctttaatt gtaagatgtc ctagacaagg aaaacctagt tacaccgtgg attggtatta
 421 ctcacaaaca aacaaaagta ttcccactca ggaaagaaat cgtgtgtttg cctcaggcca
 481 acttctgaag tttctaccag ctgcagttgc tgattctggt atttatacct gtattgtcag
 541 aagtcccaca ttcaatagga ctggatatgc gaatgtcacc atatataaaa aacaatcaga
 601 ttgcaatgtt ccagattatt tgatgtattc aacagtatct ggatcagaaa aaaattccaa
 661 aatttattgt cctaccattg acctctacaa ctggacagca cctcttgagt ggtttaagaa
 721 ttgtcaggct cttcaaggat caaggtacag ggcgcacaag tcatttttgg tcattgataa
 781 tgtgatgact gaggacgcag gtgattacac ctgtaaattt atacacaatg aaaatggagc
 841 caattatagt gtgacggcga ccaggtcctt cacggtcaag gatgagcaag cttttctct
 901 gtttccagta atcggagccc ctgcacaaaa tgaaataaag gaagtggaa ttggaaaaaa
 961 cgcaaaccta acttgctctg cttgttttgg aaaaggcact cagttcttgg ctgccgtcct
1021 gtggcagctt aatggaacaa aaattacaga ctttggtgaa ccaagaattc aacaagagga
1081 agggcaaaat caaagtttca gcaatgggct ggcttgtcta gacatggttt taagaatagc
1141 tgacgtgaag gaagaggatt tattgctgca gtacgactgt ctggccctga atttgcatgg
1201 cttgagaagg cacaccgtaa gactaagtag gaaaaatcca attgatcatc atagcatcta
1261 ctgcataatt gcagtatgta gtgtatttt aatgctaatc aatgtcctgc ttatcatcct
1321 aaaaatgttc tggattgagg ccactctgct ctggagagac atagctaaac cttacaagac
1381 taggaatgat ggaaagctct atgatgctta tgttgtctac ccacggaact acaaatccag
1441 tacagatggg gccagtcgtg tagagcactt tgttcaccag attctgcctg atgttcttga
1501 aaataaatgt ggctatacct tatgcattta tgggagagat atgctacctg gagaagatgt
1561 agtcactgca gtggaaacca acatacgaaa gagcaggcgg cacattttca tcctgacccc
1621 tcagatcact cacaataagg agtttgccta cgagcaggag gttgccctgc actgtgccct
1681 catccagaac gacgccaagg tgatacttat tgagatggag gctctgagcg agctggacat
1741 gctgcaggct gaggcgcttc aggactccct ccagcatctt atgaaagtac aggggaccat
1801 caagtggagg gaggaccaca ttgccaataa aaggtccctg aattctaaat tctggaagca
1861 cgtgaggtac caaatgcctg tgccaagcaa aattcccaga aaggcctcta gtttgactcc
1921 cttggctgcc cagaagcaat agtgcctgct gtgatgtgca aaggcatctg agtttgaagc
1981 tttcctgact tctcctagct ggcttatgcc cctgcactga agtgtgagga gcaggaatat
2041 taaagggatt caggcctc
```

IL1RL1 cDNA Sequence 2 (SEQ ID NO: 3)

NM_003856 Homo sapiens interleukin 1 receptor-like 1 (IL1RL1), transcript variant 2, mRNA

```
   1 gaggagggac ctacaaagac tggaaactat tcttagctcc gtcactgact ccaagttcat
  61 cccctctgtc tttcagtttg gttgagatat aggctactct tcccaactca gtcttgaaga
 121 gtatcaccaa ctgcctcatg tgtggtgacc ttcactgtcg tatgccagtg actcatctgg
 181 agtaatctca acaacgagtt accaatactt gctcttgatt gataaacag atggggtttt
 241 ggatcttagc aattctcaca attctcatgt attccacagc agcaaagttt agtaaacaat
 301 catggggcct ggaaaatgag gctttaattg taagatgtcc tagacaagga aaacctagtt
 361 acaccgtgga ttggtattac tcacaaacaa acaaaagtat tcccactcag gaaagaaatc
 421 gtgtgtttgc ctcaggccaa cttctgaagt ttctaccagc tgcagttgct gattctggta
 481 tttatacctg tattgtcaga gtcccacat tcaataggac tggatatgcg aatgtcacca
 541 tatataaaaa acaatcagat tgcaatgttc cagattattt gatgtattca acagtatctg
 601 gatcagaaaa aaattccaaa atttattgtc ctaccattga cctctacaac tggacagcac
 661 ctcttgagtg gtttaagaat tgtcaggctc ttcaaggatc aaggtacagg gcgcacaagt
 721 cattttggt cattgataat gtgatgactg aggacgcagg tgattacacc tgtaaattta
 781 tacacaatga aaatggagcc aattatagtg tgacggcgac caggtccttc acggtcaagg
 841 atgagcaagg cttttctctg tttccagtaa tcggagcccc tgcacaaaat gaaataaagg
 901 aagtggaaat tggaaaaaac gcaaacctaa cttgctctgc ttgttttgga aaaggcactc
 961 agttcttggc tgccgtcctg tggcagctta atggaacaaa aattacagac tttggtgaac
1021 caagaattca acaagaggaa gggcaaaatc aaagtttcag caatgggctg gcttgtctag
1081 acatggtttt aagaatagct gacgtgaagg aagaggattt attgctgcag tacgactgtc
1141 tggccctgaa tttgcatggc ttgagaaggc acaccgtaag actaagtagg aaaaatccaa
1201 gtaaggagtg tttctgagac tttgatcacc tgaactttct ctagcaagtg taagcagaat
1261 ggagtgtggt tccaagagat ccatcaagac aatgggaatg gcctgtgcca taaaatgtgc
1321 ttctcttctt cgggatgttg tttgctgtct gatctttgta gactgttcct gtttgctggg
1381 agcttctctg ctgcttaaat tgttcgtcct cccccactcc ctcctatcgt tggtttgtct
1441 agaacactca gctgcttctt ggtcatcct tgttttctaa ctttatgaac tccctctgtg
```

```
1501 tcactgtatg tgaaaggaaa tgcaccaaca accgtaaact gaacgtgttc ttttgtgctc
1561 ttttataact tgcattacat gttgtaagca tggtccgttc tatacctttt tctggtcata
1621 atgaacactc attttgttag cgagggtggt aaagtgaaca aaaaggggaa gtatcaaact
1681 actgccattt cagtgagaaa atcctaggtg ctactttata ataagacatt tgttaggcca
1741 ttcttgcatt gatataaaga aatacctgag actgggtgat ttatatgaaa agaggtttaa
1801 ttggctcaca gttctgcagg ctgtatggga agcatggcgg catctgcttc tggggacacc
1861 tcaggagctt tactcatggc agaaggcaaa gcaaaggcag gcacttcaca cagtaaaagc
1921 aggagcgaga gagaggtgcc acactgaaac agccagatct catgagaagt cactcactat
1981 tgcaaggaca gcatcaaaga gatggtgcta aaccattcat gatgaactca ccccccatgat
2041 ccaatcacct cccaccaggc tccacctcga atactgggga ttaccattca gcatgagatt
2101 tgggcaggaa cacagaccca aaccatacca cacacattat cattgttaaa ctttgtaaag
2161 tatttaaggt acatggaaca cacgggaagt ctggtagctc agcccatttc tttattgcat
2221 ctgttattca ccatgtaatt caggtaccac gtattccagg gagcctttct tggccctcag
2281 tttgcagtat acacactttc caagtactct tgtagcatcc tgtttgtatc atagcactgg
2341 tcacattgcc ttacctaaat ctgtttgaca gtctgctcaa cacgactgca agctccatga
2401 gggcagggac atcatctctt ccatctttgg gtccttagtg caatacctgg cagctagcca
2461 gtgctcagct aaatatttgt tgactgaata aatgaatgca caaccaaaaa aaaaaaaaaa
2521 aaaaaaaaaa aaaaaaaaaa aa
```

IL1RL1 cDNA Sequence 3 (SEQ ID NO: 4)

NM_173459 Homo sapiens interleukin 1 receptor-like 1 (IL1RL1), transcript variant 3, mRNA

```
   1 gaggagggac ctacaaagac tggaaactat tcttagctcc gtcactgact ccaagttcat
  61 cccctctgtc tttcagtttg gttgagatat aggctactct tcccaactca gtcttgaaga
 121 gtatcaccaa ctgcctcatg tgtggtgacc ttcactgtcg tatgccagtg actcatctgg
 181 agtaatctca acaacgagtt accaatactt gctcttgatt gataaacaga atggggtttt
 241 ggatcttagc aattctcaca attctcatgt attccacagc agcaaagttt agtaaacaat
 301 catggggcct ggaaaatgag gctttaattg taagatgtcc tagacaagga aaacctagtt
 361 acaccgtgga ttggtattac tcacaaacaa acaaaagtat tcccactcag gaaagaaatc
 421 gtgtgtttgc ctcaggccaa cttctgaagt ttctaccagc tgcagttgct gattctggta
 481 tttatacctg tattgtcaga agtcccacat tcaataggac tggatatgcg aatgtcacca
 541 tatataaaaa acaatcagat tgcaatgttc cagattattt gatgtattca acagtatctg
 601 gatcagaaaa aaattccaaa atttattgtc ctaccattga cctctacaac tggacagcac
 661 ctcttgagtg gtttaagaat tgtcaggctc ttcaaggatc aaggtacagg gcgcacaagt
 721 cattttttggt cattgataat gtgatgactg aggacgcagg tgattacacc tgtaaattta
 781 tacacaatga aaatggagcc aattatagtg tgacggcgac caggtccttc acggtcaagg
 841 tttggtgtca gagtttctgc aaattaaaaa agagcttaat ctttagtaat actcattgga
 901 ttcaaagtct aatgagaggc tttgtgatgg tatactatgg tgtacataaa tgttgtcgag
 961 tggtttttaa tctttgtttg caatactttc aacatcatca atggccttga atgagcaagg
1021 cttttctctg tttccagtaa tcggagcccc tgcacaaaat gaaataaagg aagtggaaat
1081 tggaaaaaac gcaaacctaa cttgctctgc ttgttttgga aaaggcactc agttcttggc
1141 tgccgtcctg tggcagctta atggaacaaa aattacagac tttggtgaac caagaattca
1201 acaagaggaa gggcaaaatc aaagtttcag caatgggctg gcttgtctag acatggtttt
1261 aagaatagct gacgtgaagg aagaggattt attgctgcag tacgactgtc tggccctgaa
1321 tttgcatggc ttgagaaggc acaccgtaag actaagtagg aaaaatccaa gtaaggagtg
1381 tttctgagac tttgatcacc tgaactttct ctagcaagtg taagcagaat ggagtgtggt
1441 tccaagagat ccatcaagac aatgggaatg gcctgtgcca taaaatgtgc ttctcttctt
1501 cgggatgttg tttgctgtct gatctttgta gactgttcct gtttgctggg agcttctctg
1561 ctgcttaaat tgttcgtcct ccccactcc ctcctatcgt tggtttgtct agaacactca
1621 gctgcttctt tggtcatcct tgttttctaa ctttatgaac tccctctgtg tcactgtatg
1681 tgaaaggaaa tgcaccaaca accgtaaact gaacgtgttc ttttataact
1741 tgcattacat gttgtaagca tggtccgttc tataccctttt tctggtcata atgaacactc
1801 attttgttag cgagggtggt aaagtgaaca aaaaggggaa gtatcaaact actgccattt
1861 cagtgagaaa atcctaggtg ctactttata ataagacatt tgttaggcca ttcttgcatt
1921 gatataaaga aatacctgag actgggtgat ttatatgaaa agaggtttaa ttggctcaca
1981 gttctgcagg ctgtatggga agcatggcgg catctgcttc tggggacacc tcaggagctt
2041 tactcatggc agaaggcaaa gcaaaggcag gcacttcaca cagtaaaagc aggagcgaga
2101 gagaggtgcc acactgaaac agccagatct catgagaagt cactcactat tgcaaggaca
2161 gcatcaaaga gatggtgcta aaccattcat gatgaactca ccccccatgat ccaatcacct
2221 cccaccaggc tccacctcga atactgggga ttaccattca gcatgagatt tgggcaggaa
```

```
2281 cacagaccca aaccatacca cacacattat cattgttaaa ctttgtaaag tatttaaggt
2341 acatggaaca cacgggaagt ctggtagctc agcccatttc tttattgcat ctgttattca
2401 ccatgtaatt caggtaccac gtattccagg gagcctttct tggccctcag tttgcagtat
2461 acacactttc caagtactct tgtagcatcc tgtttgtatc atagcactgg tcacattgcc
2521 ttacctaaat ctgtttgaca gtctgctcaa cacgactgca agctccatga gggcagggac
2581 atcatctctt ccatctttgg gtccttagtg caatacctgg cagctagcca gtgctcagct
2641 aaatatttgt tgactgaata aatgaatgca caaccaaaaa aaaaaaaaaa aaaaaaaaaa
2701 aaaaaaaaaa aa
```

[0252] Following are show human amino acid sequences for isoform 1, isoform 2 and isoform 3 of
IL1RL1.

IL1RL1 Amino Acid Sequence 1 (SEQ ID NO: 5)

NP_057316 interleukin 1 receptor-like 1 isoform 1 precursor; interleukin 1 receptor-related protein;
homolog of mouse growth stimulation-expressed gene; ST2 protein [Homo sapiens]

MGFWILAILTILMYSTAAKFSKQSWGLENEALIVRCPRQGKPSYTVDWYYSQTNKSIPT
QERNRVFASGQLLKFLPAAVADSGIYTCIVRSPTFNRTGYANVTIYKKQSDCNVPDYLM
YSTVSGSEKNSKIYCPTIDLYNWTAPLEWFKNCQALQGSRYRAHKSFLVIDNVMTEDA
GDYTCKFIHNENGANYSVTATRSFTVKDEQGFSLFPVIGAPAQNEIKEVEIGKNANLTCS
ACFGKGTQFLAAVLWQLNGTKITDFGEPRIQQEEGQNQSFSNGLACLDMVLRIADVKE
EDLLLQYDCLALNLHGLRRHTVRLSRKNPIDHHSIYCIIAVCSVFLMLINVLVIILKMFWI
EATLLWRDIAKPYKTRNDGKLYDAYVVYPRNYKSSTDGASRVEHFVHQILPDVLENKC
GYTLCIYGRDMLPGEDVVTAVETNIRKSRRHIFILTPQITHNKEFAYEQEVALHCALIQN
DAKVILIEMEALSELDMLQAEALQDSLQHLMKVQGTIKWREDHIANKRSLNSKFWKHV
RYQMPVPSKIPRKASSLTPLAAQKQ

IL1RL1 Amino Acid Sequence 2 (SEQ ID NO: 6)

NP_003847 interleukin 1 receptor-like 1 isoform 2 precursor; interleukin 1 receptor-related protein;
homolog of mouse growth stimulation-expressed gene; ST2 protein [Homo sapiens]

MGFWILAILTILMYSTAAKFSKQSWGLENEALIVRCPRQGKPSYTVDWYYSQTNKSIPT
QERNRVFASGQLLKFLPAAVADSGIYTCIVRSPTFNRTGYANVTIYKKQSDCNVPDYLM
YSTVSGSEKNSKIYCPTIDLYNWTAPLEWFKNCQALQGSRYRAHKSFLVIDNVMTEDA
GDYTCKFIHNENGANYSVTATRSFTVKDEQGFSLFPVIGAPAQNEIKEVEIGKNANLTCS
ACFGKGTQFLAAVLWQLNGTKITDFGEPRIQQEEGQNQSFSNGLACLDMVLRIADVKE
EDLLLQYDCLALNLHGLRRHTVRLSRKNPSKECF

IL1RL1 Amino Acid Sequence 3 (SEQ ID NO: 7)

NP_775661 interleukin 1 receptor-like 1 isoform 3 precursor; interleukin 1 receptor-related protein;
homolog of mouse growth stimulation-expressed gene; ST2 protein [Homo sapiens]

MGFWILAILTILMYSTAAKFSKQSWGLENEALIVRCPRQGKPSYTVDWYYSQTNKSIPT
QERNRVFASGQLLKFLPAAVADSGIYTCIVRSPTFNRTGYANVTIYKKQSDCNVPDYLM
YSTVSGSEKNSKIYCPTIDLYNWTAPLEWFKNCQALQGSRYRAHKSFLVIDNVMTEDA

GDYTCKFIHNENGANYSVTATRSFTVKVWCQSFCKLKKSLIFSNTHWIQSLMRGFVMV
YYGVHKCCRVVFNLCLQYFQHHQWP

[0253] Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the aspects which follow. All publications or patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[0254] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents. U.S. patents and other publications referenced herein are hereby incorporated by reference.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No.	EL 961008095 US

INVENTOR(S)

Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Steven	MAH	San Diego, California
Andreas	BRAUN	San Diego, California
Stefan M.	KAMMERER	San Diego, California

Additional inventors are being named on the __1__ *separately numbered sheets attached hereto*

TITLE OF THE INVENTION (500 characters max)

METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF

CORRESPONDENCE ADDRESS

Direct all correspondence to:

[X] Customer Number:	25225

OR

[ ] Firm or Individual Name	
Address	

City		State		Zip	
Country		Telephone		Fax	

ENCLOSED APPLICATION PARTS (check all that apply)

[X] Specification *Number of Pages*	121	[ ] CD(s), Number		
[X] Drawing(s) *Number of Sheets*	3	[X] Other	Return Receipt Postcard	
[X] Application Data Sheet. See 37 CFR 1.76 (4 pages)		(specify):		

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

		FILING FEE AMOUNT ($)
[X] Applicant claims small entity status. See 37 CFR 1.27.		
[ ] A check or money order is enclosed to cover the filing fees.		
[X] The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 03-1952		80.00
[ ] Payment by credit card. Form PTO-2038 is attached.		

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

[X] No   [ ] Yes, the name of the U.S. Government agency and the Government contract number are: _____

[Page 1 of 2]

Respectfully submitted,

SIGNATURE		Date	April 1, 2004
TYPED OR PRINTED NAME	Bruce D. Grant	REGISTRATION NO. *(if appropriate)*	47,608
TELEPHONE	(858) 720-7962	Docket Number:	524593009000

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

SD-189456

PROVISIONAL APPLICATION COVER SHEET
Additional Page

Docket Number	524593009000

INVENTOR(S)/APPLICANT(S)		
Given Name (first and middle [if any])	Family or Surname	Residence (City and either State or Foreign Country)
Matthew Roberts	NELSON	San Marcos, California
Rikard Henry	RENELAND	San Diego, California
Maria L.	LANGDOWN	San Diego, California

sd-189456